PB80-181894

Pilot Study - Uses of Medicare Morbidity Data in Health Effects Research

System Sciences, Inc.
Chapel Hill, NC

Prepared for

Health Effects Research Lab.
Research Triangle Park, NC

Mar 80

U.S. Department of Commerce
National Technical Information Service
NTIS

United States
Environmental Protection
Agency

Health Effects Research
Laboratory
Research Triangle Park NC 27711

EPA-600/1-80-018
March 1980

PB80-181894

Research and Development

EPA

Pilot Study—Uses of Medicare Morbidity Data in Health Effects Research

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U. S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.
EPA-600/1-80-018

2.

3. RECIPIENT'S ACCESSION NO.
PB 80 181894

4. TITLE AND SUBTITLE
Pilot Study - Uses of Medicare Morbidity Data In Health Effects Research

5. REPORT DATE
March 1980

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
Irwin J Shiffer & Edgar A Parsons

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
System Sciences, Inc.
P.O.Box 2345
Chapel Hill, NC 27514

10. PROGRAM ELEMENT NO.
1AA817

11. CONTRACT/GRANT NO.
68-02-2782

12. SPONSORING AGENCY NAME AND ADDRESS
Health Effects Research Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, NC 27711

13. TYPE OF REPORT AND PERIOD COVERED

14. SPONSORING AGENCY CODE
EPA 600/11

15. SUPPLEMENTARY NOTES

16. ABSTRACT

This project is a pilot investigation of the practicability of utilizing Social Security Administration (SSA) Medicare morbidity data to supplement mortality data in cancer and other environmentally-related studies. For this study non-confidential data on 1.2 million hospitalizations for 815,000 persons diagnosed as having a neoplasm, respiratory, or digestive disease during 1971, 1972, and1973 were included. The data are kept current by SSA for their analysis purposes.

The Medicare files are the only known source that incorporates all cancer cases systematically, regardless of whether the cancer is fatal. The Files facilitate analysis of possible relationships between emissions of a specific industry and the disease(s) rates for the co-located population.

Numerous recommendations are made for applying the Medicare files to additional analyses of environment-to-health relationships. Prominent among these are cancer hospitalization trends, by county, augmented with cancer mortality trends, and available emmissions/monitoring measures to identify areas of cancer increase/decrease possibly related to environmental influences.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Medicare files Cancer Environment/health relationship Emmissions		06F,T

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC	UNCLASSIFIED	
	20. SECURITY CLASS *(This page)*	22. PRICE
	UNCLASSIFIED	

EPA Form 2220-1 (Rev. 4-77) PREVIOUS EDITION IS OBSOLETE

EPA-600/1-80-018
March 1980

PILOT STUDY--USES OF MEDICARE MORBIDITY DATA IN HEALTH EFFECTS RESEARCH

by

Irwin J. Shiffer and Edgar A. Parsons
System Science, Inc.
P. O. Box 2345
Chapel Hill, North Carolina 27514

Contract No. 68-02-2782

Project Officer: Wilson B. Riggan

Health Effects Research Laboratory
U.S. Environmental Protection Agency
Research Triangle Park, North Carolina 27711

U.S. ENVIRONMENTAL PROTECTION AGENCY
OFFICE OF RESEARCH AND DEVELOPMENT
HEALTH EFFECTS RESEARCH LABORATORY
RESEARCH TRIANGLE PARK, NORTH CAROLINA 27711

DISCLAIMER

This report has been reviewed by the Health Effects Research Laboratory, U.S. Environmental Protection Agency, and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement of recommendation for use.

FOREWORD

The many benefits of our modern, developing, industrial society are accompanied by certain hazards. Careful assessment of the relative risk of existing and new man-made environmental hazards is necessary for the establishment of sound regulatory policy. These regulations serve to enhance the quality of our environment in order to promote the public health and welfare and the productive capacity of our Nation's population.

The Health Effects Research Laboratory, Research Triangle Park, conducts a coordinated environmental health research in toxicology, epidemiology, and clinical studies using human volunteer subjects. These studies address problems in air pollution, non-ionizing radiation, environmental carcinogenesis and the toxicology of pesticides as well as other chemical pollutants. The Laboratory participates in the development and revision of air quality criteria documents on pollutants for which national ambient air quality standards exist or are proposed, provides the data for registration of new pesticides or proposed suspension of those already in use, conducts research on hazardous and toxic materials, and is primarily responsible for providing the health basis for non-ionizing radiation standards. Direct support to the regulatory function of the Agency is provided in the form of expert testimony and preparation of affidavits as well as expert advice to the Administrator to assure the adequacy of health care and surveillance of persons having suffered imminent and substantial endangerment of their health.

Collection of epidemiological health data is a costly and time consuming effort. For these reasons, we are always looking for sources of health data which may be used in our environmental health studies. This report describes a pilot investigation of the practicality of using Social Security Administration (SSA) Medicare 20 percent sample of short-stay hospital discharge data to supplement mortality data in cancer and other environmentally related health studies.

F. G. Hueter, Ph.D.
Director
Health Effects Research Laboratory

PREFACE

Little is known of the relationship between the complex of chemical and physical factors comprising the "real life" human environment and the effects of one or more of these factors on human health. Much of what is known is derived from laboratory animal observations and accidents or unusual occurrences involving humans. While helpful, the non-human data and the atypical events are generally inadequate from the perspective of relevance to human environmental exposure parameters. Serious technical questions can also be posed concerning the instrumentation measuring the composition of the air, and whether fixed/mobile monitoring data can be re-assembled to characterize accurately the air actually breathed by humans of different ages and living patterns.

A supplementary approach to determining factors harmful to human health in the "real life" environment is to stress geographical areas and/or selected populations characterized by unusually high/low levels of indicator diseases, such as respiratory diseases and cancer. The basic objective of EPA being human health protection, this approach offers the advantages of identifying areas of greatest need, and of enhancing responsiveness to reduce harmful environmental pollutants. An additional advantage is deferment/avoidance of costly abatement and control measures in areas of lesser need (or no need) without jeopardizing health.

This approach requires health status details rarely available because of cost and confidentiality barriers. Considerable EPA reliance has been, therefore, placed on the low-cost non-confidential mortality statistics collected and stored on computer tape by the National Center for Health Statistics for purposes that predate EPA. These tapes exclude much more information of interest to EPA which is collected in hard copy and presents considerable cost and clearance problems. Serious attention has not heretofore been directed by EPA to utilization of other established local-state-Federal health information acquisition systems featuring non-mortality data.

In accordance with its unsolicited proposal that provided the basis for this project, System Sciences, Inc. staff continued explorations with respect to environmental health applications of the "Medicare morbidity" files. These files were created by the Social Security Administration and are being updated annually for non-environmental purposes. SSA cooperation was outstanding. Staff of the Program Statistics Division, Office of Research and Statistics, and Mr. Charles Fisher in particular, immediately perceived the environmental health relevance of certain data elements. Tapes were prepared and transferred to EPA that met confidentialy constraints with extracts of most of the data pertinent to health conditions and the environment. The tapes and data proved satisfactory for pilot investigation purposes.

The underlying goal of this project is to demonstrate the applicability to EPA of the partial Medicare files that have been obtained. Assuming applicability, additional SSA information may be provided subject always to confidentiality precautionary requirements. Properly used in conjunction with emissions, monitoring data, demography, mortality and other "computerized" EPA resources, the Medicare files have been demonstrated to offer unprecedented opportunities for contributing to knowledge of relationships between the environment and human health.

ABSTRACT

This project is a pilot investigation of the practicability of utilizing Social Security Administration (SSA) Medicare morbidity data to supplement mortality data in cancer and other environmentally-related studies. The Medicare files are a 20 percent random selection, aggregating 5,000,000 of the 25,000,000 insurees; for this study non-confidential data on 1.2 million hospitalizations for 815,000 persons diagnosed as having a neoplasm, respiratory, or digestive disease during 1971, 1972, and 1973 were included. The data are kept current by SSA for their analysis purposes.

Project outputs demonstrate the availability to EPA of a major new data source for health effects and program formulation analyses. This project prepared the Medicare data for use directly with emissions, ambient air, demography, mortality, and other data banks at EPA's National Computer Center. Outputs include tables and computer-prepared multi-colored maps illustrating relationships between any selected industry of interest and respiratory or other diseases of the co-located population.

The Medicare codes show hospitalization date, diagnosis, age, race, sex and county and state of residence. When used to measure the Pittsburgh 1973 air pollution episode, the files disclosed a significant increase in respiratory disease hospitalizations immediately prior to, and during the episode. With an annual average of 16 hospitalizations to each death, the Medicare files increase substantially the opportunity for identifying possible relationships between particular environments and the co-located populations. Migration aspects can be included as appropriate.

The Medicare files are the only known source that incorporates all cancer cases systematically, regardless of whether the cancer is fatal. Of new 1979 cancer cases (765,000), about 52 percent (395,000) result in cancer death. Cancer mortality files exclude, by definition, the non-fatal neoplasms not

meeting the "underlying cause of death" definition as well as deaths from accidents, and greatly understate cancer prevalence. Project work included development of a procedure using morbidity and mortality data together for maintaining, by county, a "running inventory" of increases/decreases in cancer "survivors," with recommendations for comparative analyses with emissions and ambient data.

The Medicare files facilitate analysis of possible relationships between emissions of a specific industry and the disease(s) rates for the co-located population. In demonstrating the use of these files in response to a special EPA request, it was found that average Medicare hospitalization rates for each of six respiratory disease classifications are lower for the nation's 97 steel-producting counties than for the non-steel-producing counties. Similarly, the 41 steel-producing counties of Ohio and Pennsylvania also have lower respiratory disease hospitalizations than the non-steel counties in these two states. This finding was unexpected. Without further investigation, it should not be concluded from this demonstration that steel-producing industry emissions are less harmful than emissions of other industries, or that similar conclusions are justified. A definitive analysis was not within the scope of this project.

Numerous recommendations are made for applying the Medicare files to additional analyses of environment-to-health relationships. Prominent among these are cancer hospitalization trends, by county, augmented with cancer mortality trends, and available emissions/monitoring measures to identify areas of cancer increase/decrease possibly related to environmental influences.

Millions of health status files have been demonstrated to be available for EPA use, and essentially without cost for data collection or computer encoding.

CONTENTS

Foreword iii
Preface iv
Abstract vi
Figures x
Tables xi
Abbreviations xiv

I. Introduction 1
II. Conclusions 3
III. Recommendations 11
  A. Applications 12
  B. Supplementary Files-Extensions 18
  C. Improvements 20
IV. Data Resources, Methods, and Findings 21
  A. Data Resources 21
  B. Methods and Findings 27
V. Review and Discussion 101
  A. Berks County--Morbidity/Mortality Comparison by Year 102
  B. Daily Morbidity/Mortality During Pittsburgh Episode of 1973 104
  C. Morbidity in Steel-Producing Areas 106
  D. Migration of Medicare Claimants 108
  E. Findings and Conclusions 109

FIGURES

Number		Page
IV-1	U.S. counties with low migration for white males of all ages	39
IV-2	U.S. counties with low migration of white males over age of 44	40
IV-3	High in-migration counties of interest in Florida . . .	45
IV-4	High in-migration county of interest in New Jersey . . .	46
IV-5	Counties of interest in Pennsylvania	47
IV-6A	Levels of employment in two Florida counties for the years 1953, 1959, 1965, 1970 and 1974	52
IV-6B	Levels of employment in five Pennsylvania counties for the years 1953, 1959, 1965, 1970 and 1974	53
IV-7	U.S. counties in which are located manufacturing establishments with SIC Code 3312	86
IV-8	Steel-producing counties categorized by Medicare respiratory disease hospital patients over 64 years of age, 1971-1973	97
IV-9	Ohio and Pennsylvania counties classified by steel industry suspended particulate emissions and Medicare claimant respiratory disease hospital rates	98

TABLES

Number		Page
III-1	Type of Sampling, Pollutant, and Time Period of Data for Birmingham, Alabama	16
IV-1	Record Layout for Data Tapes from Medicare Hospital Discharge Survey	23
IV-2	Counties with Highest White Male Mortality Rates for all Neoplasms	25
IV-3	County Business Patterns--Florida (Pinellas County) . . .	26
IV-4	Sample Point Source Listings from National Emissions Data System for Particulate and Sulfur Oxides Emissions (in Short Tons, 1977)	28
IV-5	Sample Display of Medicare Data Assembled Longitudinally to Show Hospitalization Causes, 1971-1973	30
IV-6	Beneficiary Identification Codes Social Security Administration Type of Claim	31
IV-7	Beneficiary Identification Code Railroad Board Employee Type of Claim	32
IV-8	Patient Non-Medical Characteristics from Consolidated Medicare Records for 1971-1973	33
IV-9	Disease Prevalence from Consolidated Medicare Records for 1971-1973	34
IV-10	Disease Category Combinations from Consolidated Medicare Records for 1971-1973	35
IV-11	U.S. Counties with Total Population Greater than 10,000 in Which the Ratio of In-Migration to Out-Migration for White Males Who Were 45 Years of Age or Older in 1970 is Greater Than 3.00	41
IV-12	U.S. Counties with Total Population Greater than 10,000 in Which the Ratio of Out-Migration to In-Migration for White Males Who Were 45 Years of Age or Older in 1970 is Greater Than 2.00	43
IV-13	County Population Characteristics	48
IV-14	Medicare Data Characteristics	49
IV-15	County Urban/Industrial Characteristics	50

TABLES (CONTINUED)

Number | Page

IV-16 Death Counts by Neoplasm Type and Age Group for Berks County, PA and Years 1971, 1972, and 1973 58

IV-17 Death Counts by Neoplasm Type for Persons Aged 65 or Over in Berks County, PA 59

IV-18 Use of Medicare Data to Estimate Cancer Prevalence . . . 60

IV-19 Allegheny County, PA, Residents Who Died in the County, August 1 Through September 22, 1971-1973, with Subtotals for Underlying Death Cause of Respiratory or Cardiovascular Disease 62

IV-20 Daily Admissions by Disease Class in Allegheny County, PA, During the Period August 13 Through September 20, 1971 64

IV-21 Daily Admissions by Disease Class in Allegheny County, PA, During the Period August 13 Through September 20, 1972 65

IV-22 Daily Admissions by Disease Class in Allegheny County, PA, During the Period August 13 Through September 20, 1973 66

IV-23 Daily Counts of Allegheny County First Admissions During August 13-September 10, 1971, 1972, 1973 by Respiratory Disease Category 67

IV-24 Numbers of Persons Admitted with Indicated Diagnoses Each 13-Day Period: August 13-25, August 26-Sept. 7, and Sept. 8-Sept. 20, 1971-1973 68

IV-25 Daily Admissions for Pneumonia and Bronchitis, Emphysema, or Asthma in the Pittsburgh SMSA During the 39-Day Period Including the 1973 Air Pollution Episode and Comparable Period in 1971 and 1972 71

IV-26 Illustrative Distributions of Time Spans from First Admission for Cancer to Last Admission, 1971-1973 . . . 74

IV-27 Average Months and Rank from First Medicare Cancer to Death of Various Causes 76

IV-28 Calculation of Median Survival Times for Cancer Patients Age 65 and Older, and Comparison of Relative NCI and Medicare Rankings 78

IV-29 Frequency Distribution of Numbers of (A) Counties of Residence, (B) States of Residence, and (C) Counties of Residence within Indicated State During 1971-1973 for Medicare Claimants Residing in Indicated State and County 80

TABLES (CONTINUED)

Number Page

IV-30 List of 31 States, 97 Counties, Two Independent Cities of the United States in Which are Located Manufacturing Establishments with SIC Code 3312 85

IV-31 Data Entry Form 87

IV-32 Total Number of Medicare Claimants (with Corresponding Morbidity Rates) in 97 Steel-Producing Counties and All U.S. Counties for Hospitalizations in 1971-1973 with Diagnoses in Six Respiratory Disease Categories . . 89

IV-33 Respiratory Disease Categories for Which Medicare Claimant Rates in "Steel-Producing" Counties are Greater than Respective State Rates 91

LIST OF ABBREVIATIONS

CO: Carbon monoxide.

EPA: Environmental Protection Agency.

HERL: Health Effects Research Laboratory.

HEW: Department of Health, Education, and Welfare.

ICDA: International Classification of Diseases, Abstracted.

NEDS: National Emissions Data System.

NO_x: Oxides of nitrogen.

OAQPS: Office of Air Quality Planning and Standards.

RAMS: Regional Air Monitoring System.

RAPS: Regional Air Pollution Study.

RSP: Respirable suspended particulates.

SAROAD: Storage and Retrieval of Aerometric Data System.

SMSA: Standard Metropolitan Statistical Area.

SO_2: Sulfur dioxide.

SSA: Social Security Administration.

SSI: System Sciences, Inc.

TSP: Total suspended particulates.

SECTION I

INTRODUCTION

This document describes the work and findings of a pilot study investigating the potential utility for epidemiologic research of data from the Social Security Administration's Medicare short-stay hospital discharge sample survey. The work was performed by System Sciences, Inc. under the direction of the Health Effects Research Laboratory of the Environmental Protection Agency in fulfillment of EPA Contract No. 68-02-2782. Certain printout reproductions, tables, maps, graphs, and data source descriptive materials are included to meet contractual report requirements for such explanatory materials related to the various project work steps.

A prime consideration was the development of systematic procedures for relating morbidity data from this new resource to county level mortality data, industrial data, environmental data, and demographic data. Many of the procedures utilized the UNIVAC 1110 at EPA's National Computer Center, Research Triangle Park, N.C. The procedures facilitate analyses of a single county, a large number of counties, Standard Metropolitan Statistical Areas, states, and national trends.

For pilot study purposes, particularly constraints of economy, timeliness, highly confidential files, and the practical risks of unforeseeable obstacles in applying Medicare medical histories for unprecedented purposes, only three categories of hospital diagnoses were investigated. Any data that might jeopardize confidentiality were deleted; only the minimum facts essential to the pilot study were extracted from the Medicare tapes. The categories selected are of special interest to the Environmental Protection Agency--all respiratory diseases, neoplasms, and diseases of the digestive system. The Medicare tapes for these three disease categories yielded about 1.2 million hospitalizations, for some 815,000 individuals. Disease definitions and codes as provided by the Social Security Administration are in conformance with Eighth Edition, International Classification of Diseases, Abstracted (ICDA).

The study was performed in the following steps. First, the computer tapes with selected data fields of the Medicare files were verified and reformatted for computer-assisted analyses to fulfill study requirements. Then, counts were made to determine the demographic and other characteristics of the sample file. Next, population, migration, environmental, and industrial data were accessed and studied to select areas for further study. Employment histories were obtained for the selected counties to portray industrial development patterns in detail.

Three in-depth investigations were performed to demonstrate morbidity/mortality comparisons: (1) annual morbidity data for one county were examined to determine the feasibility of obtaining, for environmental health indicator purposes, rates of death due to selected diseases, and concurrently the accumulating "inventory" of disease experience, which may or may not eventually cause death in the surviving population of an area; (2) daily admissions during a period centered on an air pollution episode were studied; and (3) cancer patient survival times from the Medicare data were compared with National Cancer Institute times as a partial test of Medicare data reliability.

Another investigation was directed to the use of Medicare patient state/county residence at times of treatment to assess the impact of migration on epidemiological studies.

Finally, to demonstrate the utility of the Medicare morbidity data base for environmental health research, an in-depth analysis was performed comparing counties with steel-production facilities to other counties and to the nation as a whole. For counties in Ohio and Pennsylvania, the comparison utilized suspended particulate emissions as the measure of the relative pollution significance of steel industry sources to all "point" sources in each county. Respiratory diseases were selected as a related measure of health effects.

SECTION II

CONCLUSIONS

This chapter summarizes findings and conclusions derived from the pilot study investigating the potential utility for epidemiological research of the Medicare hospitalization files. Priority and emphasis was given to the feasibility of applications on behalf of responsibilities of the Environmental Protection Agency (EPA), particularly airborne pollution of industrial origin.

Enough has been learned in this pilot investigation to conclude that it is feasible to apply the data bases described herein to make major contributions to environmental health knowledge by identifying relationships (if any) between specific industries, emissions, air quality, and selected diseases indicative of human health status.

The different data bases can also be utilized to determine trends in environmental health benefits (if any) from air pollution abatement programs of the recent past--for example, since 1970.

Findings from the preliminary investigation comparing steel- and non-steel-producing counties by rates of Medicare respiratory disease hospitalizations suggest the strong possibility of erroneous and misleading assumptions concerning co-locational relationships between certain industries, emissions, and the available human health evidence.

Little is now known of human health benefits from environmental programs, although Governmental actions are essentially justified on grounds of protecting or improving the public health. Government and private industrial costs can be estimated with some reliability. Benefits are actually measured in terms of chemical content decreases, with little or no means of estimating the human health benefits from the decrease. With the Medicare data, comparisons of carefully selected areas, with and without improvements in air quality, by rates and causes of selected hospitalization diagnoses, will shed light on the prime feature of EPA programs, and about which least is known--the human health benefits. The next chapter deals with recommendations for further work in this area.

The following points are believed especially germane to responsibilities of Health Effects Research Laboratory, Office of Research and Development, EPA.

1. Medicare file contents are credible and reliable descriptors.

2. Privacy and confidentiality constraints can be respected without serious detriment to the full value of the file.

3. The 20 percent sample selection process is random, with selectees aggregating some 5,000,000 persons of the 25,000,000 eligible, and provides coordinated local, state, and national representativeness.

4. Because there are over 16 times more hospitalizations than deaths, the files multiply environmental health perceptiveness, and provide unique environmental health data not available from mortality files. For examples, the Medicare files provide the only county-state-national coordinated source for the following:

   a. Non-fatal diseases of environmental health interests, and
   b. Potentially fatal cases of environmental health interest where death intervenes from causes of lower environmental health interest (accidents, strokes to lung cancer patients, etc.).

5. Time-phased histories from first Medicare hospitalization to death, regardless of changes in patient residence, are routinely computer-encoded. This capability is particularly relevant to cancer or other long latency diseases.

6. The hospitalization and mortality files have distinct advantages, and should be used in mutual support, not exclusion, of one another.

7. The files may be used directly to evaluate the environmental health status of the country generally, or to compare one area with another with the precision of any codifiable hospitalization diagnosis, by

   a. absolute numbers of selected hospitalization diagnoses
   b. rates per X population of selected hospitalization diagnoses
   c. trends in absolute numbers
   d. trends in rates per X population
   e. increasing/decreasing surviving population with selected hospitalization diagnoses
   f. increasing/decreasing rates of surviving X population with selected hospitalization diagnoses.

8. The Medicare hospitalization files can be used to quantify health effects of unusual temporary environmental conditions, such as an air pollution episode. (Hospitalization admission dates are routinely computer encoded.)

9. The Medicare hospitalization files offer a unique capability for relating admissions/rates of selected diagnoses to the chemical composition of the air. This capability is believed particularly valuable for specific geopolitical areas with the most complete and reliable data, such as St. Louis.

10. Numbers and rates of hospitalizations from environmentally-related diseases can be used to assess the co-locational relationship to industries of major environmental pollution-concern. The presence/absence of significant co-locational health relationship has agency-wide regulatory and administrative implications, and should be expanded to include assessments of emissions, ambient data, pollutants from additional point and area sources, and other factors as appropriate.

11. The Medicare hospitalization files are in tape format, readily transferrable to EPA's UNIVAC 1110 at Research Triangle Park, and complexities

of county coding differences, longitudinal assemblage of hospitalizations from annual tapes, and compatibility with EPA SAROAD, NEDS and other major data bases have been demonstrated to be relatively minor. Special note is made of the ease and low cost with which multi-color computer-prepared maps may be made for the United States with detail to the county level. These maps are helpful in presenting not only the "big picture" but also in facilitating the analysis of possible interrelationships between industry, emissions, and selected diseases in the actual environment.

12. The full value to EPA of the Medicare files cannot yet be assessed, as many possibilities have not been examined. For example, the classification system enables differentiation of hospitalization rates by sex and race between patients who are "primary beneficiaries" or employees, and those who are Medicare enrollees because of their relationship to the primary beneficiary. Conceptually, one might thus categorize certain diseases as environmental or occupational.

The general conclusions of this pilot study are presented below, together with an abbreviated description of the basis for the conclusions. For details, the reader is referred to the text, particularly Section IV, Data Resources, Methods, and Findings.

1. Credibility--The Medicare hospitalization file extracts made available to EPA and SSI offer credible and reliable health indicators of a population of particular environmental interest--the elderly, primarily age 65 and over, aggregating over 25 million, and with 20 percent of these constituting the Medicare file made available.

   Basis--Each of ten neoplasms, selected for definition uniformity with National Cancer Institute records, were checked for "survival times" (time to death from confirmed diagnosis), and death percentages attributed to the neoplasm. There was close correspondence for each of the ten neoplasm types.

2. National and Local Representativeness--The Medicare sample is large, and the Medicare population is distributed throughout the country in such manner that the Medicare tapes provide the potential for continuous monitoring of this particularly sensitive population through selected, environmentally-related diseases. This has many applications to EPA responsibilities for evaluation of pollution sources and the health of co-located populations. No purpose is perceived that would justify increasing the 20 percent sample, except for special, localized, intensive ad hoc investigations.

   Basis--Each hospitalization is coded by discharge diagnosis, date of admission, county/state of residence, age, race, sex, and the 5,000,000 persons in the sample reside in every county in the United States. The 20 percent selection process is random.

3. Longitudinal Health Analyses--The Medicare files can be assembled to show, longitudinally, the Medicare hospitalization history from first admission to death for each Medicare sample selectee, regardless of residence changes. Assuming continued cooperation with HEW, EPA thus has a systematic procedure for obtaining knowledge of changes in health conditions over time, and non-fatal disease occurrences, by county of residence at time of hospitalization.

   Basis--See example printout presented in text.

4. Accumulative Disease Capabilities Facilitates Environmental Effects Comparisons--The Medicare files can be assembled to compute the population surviving from one or more hospitalizations of selected environmentally-related diseases. Accumulations of patients by disease may be by county, counties, or with special permission, by Zip Codes. There are ten times more Zip Codes than counties. This capability is unique. EPA applications extend from determining relationship (if any) between pollution sources and selected diseases, to determining priorities for abatement and control.

Basis--See example calculations for Berks County, Pennsylvania.

5. **Change in Rates and Trends of the Surviving Population with Specific Hospitalization Causes Offers a Powerful EPA Health Analysis Tool with Particular Applications to Cancer and the Environment**--Increases/decreases of the Medicare population surviving selected environmentally-related disease hospitalization can be assembled in such a manner to show "small area" trends, and accumulative status. Conceptually, the increases/decreases may provide an index of success/failure for EPA programs/projects justified on environmental health improvement grounds.

   Basis--See Berks County, Pennsylvania, example of net annual increase of surviving cancer cases, by type, and cumulative prevalence. (Any defined and encoded disease(s) may be selected.)

6. **Medicare Hospitalization Files Appear Ideally Suited to Quantifying Health Effects of Unusual Environmental Conditions**--Medicare hospitalizations offer EPA significant potential for measuring promptly the health consequences of adverse environmental circumstances, such as an air pollution episode. Conceptually, the ambient and hospitalization measurements could be concurrent. Formal/legal emergency episode determinations, and countermeasures by Federal, State, and local environmental and public health authorities would be greatly assisted by an experienced-based air quality-health linkage system.

   Basis--The 1973 Allegheny County August 26 - September 7 air pollution episode, and the days immediately prior thereto, were noteworthy because of a significant increase in Medicare respiratory disease hospitalizations. Similarly, more mortalities from respiratory diseases and cardiovascular diseases were recorded for people of all ages during the episode than prior to or following the episode. See tables in text.

7. **Changes in residence of Medicare hospitalization claimants are unusual.**

Basis--Among Medicare hospitalization claimants, 1971 through 1973, in two Florida and five Pennsylvania counties, there were insignificant changes in state of residence, and less than five percent residence relocations among counties. Essentially all of the relocations were within Florida.

8. The presence of a steel-producing facility does not necessarily cause higher Medicare respiratory disease rates.

   Basis--Average Medicare hospitalization rates for each of six respiratory disease classifications are lower for the 97 U.S. steel-producing counties than for the non-steel-producing counties. Respiratory disease rates for white males are significantly higher than for white females. These facts are based on computer-assisted analysis of approximately 280,000 Medicare respiratory disease patients--see text and tables.

9. The relative quantities of TSP from steel-producing sources in a county do not necessarily cause high Medicare respiratory disease rates.

   Basis--The 20 Ohio and Pennsylvania counties with the lowest percentage of steel industry TSP emissions, and the 21 counties with the highest percentage of TSP emissions, showed an average for each of six categories of Medicare respiratory disease hospitalizations that was lower than the average rate for the 104 counties in these two states that did not contain steel-producing facilities.

10. Analyses of relationships between carcinogenic hazardous industries, mortalities, and Medicare neoplasm hospitalizations must include migration and other factors to avoid misleading conclusions. It is also necessary to include more than three years of Medicare hospitalization data to exploit fully the Medicare file source potential for contributing to knowledge of environmental sources of cancer.

Basis--The Florida counties had high rates of elderly in-migration, essentially no carcinogenic hazardous industries, and compared with the Pennsylvania counties studied, higher lung cancer rates, and all cancer rates. The Pennsylvania counties had a comparatively stable population, and some chemical, petroleum, rubber, and primary steel industry employment. Non-Florida exposures to environmental pollution are believed to account for much of the higher Florida cancer rates.

For the counties studied, no significant carcinogenic industry source was identified as introduced within a time span that would affect the 1971-1973 hospitalizations. To detect the impact of a newly-introduced carcinogenic source, it is necessary to have appropriate baseline experience. Three years of data are insufficient for both baseline and impact measurement of diseases with long latency periods. For hospitalizations caused or exacerbated by acute temporary changes in the environment, such as an emergency episode, the three years of data enable many unanswered questions to be resolved.

SECTION III

RECOMMENDATIONS

The Medicare hospitalizations file has been demonstrated in this pilot investigation to have extensive potential for determining relationships between the environment and human health. The contents of this constantly growing file offer the only single source documenting consistently the extent of fatal and non-fatal diseases related to the environment. From a health policy perspective, the contents offer the possibility of determining the human health improvements resulting from the billions of dollars expended, and planned for expenditure. From a technical perspective, the Medicare file contents may also provide guidance as to the human health benefits, if any, resulting from different control and abatement technologies as applied to improve air quality.

The purpose of this section is to present recommendations responsive to the project's scope of work requirement for follow-up steps appropriate to the findings and conclusions of the pilot investigation. The recommendations are to include, but are not limited to, use of the Medicare hospitalization data for determining carcinogenic health effects as early as feasible. Recommendations are presented under three broad categories, as follows:

(1) Applications--Apply the Medicare hospitalizations file to analyses/products of current HERL and EPA interest;

(2) Supplementary files, new data, and pilot investigations--Extend and/or expand the Medicare pilot investigation files to obtain knowledge of different relationships between the non-industrial or background environment and human health, between different categories of urbanized-industrial areas and the health of the colocated

population, and explore the feasibility of alternative techniques for utilizing supplementary data files collected for non-EPA purposes on behalf of HERL and EPA responsibilities; and

(3) Improvements--Verify, re-format, and otherwise improve the reliability, accuracy, and utility of the 1971-1973 computerized Medicare files presently in hand, and others on order.

The recommendations listed for follow-up to this pilot investigation do not purport to be complete. Several are closely related to one another. Collectively, the recommendations present extensive and intensive steps now feasible to be taken by EPA in bridging the knowledge gap between evidence of human health and the measureable components of the air being regulated, and controlled by EPA.

In preparing these recommendations, it has been assumed that readers have some technical familiarity with the contents and limitations of EPA information-acquisition systems. An understanding of certain recommendations will be helped by a review of the following chapter, which describes the data resources utilized in this pilot investigation. Other recommendations are based on the demonstrated credibility and reliability of the Medicare hospitalizations data, as presented briefly in the preceding section. These additional recommendations are directed to applications not attempted or not considered seriously in the absence of the pilot investigation findings.

A. APPLICATIONS

1. Investigate Alternative Methodologies for Identifying Areas of Increasing Cancer Prevalence and Recommend a Methodology for Use by EPA

Integrate cancer mortality and Medicare cancer hospitalization files for selected areas, and feature trend analyses in the investigation and comparison. (Note: The widely publicized NCI "Cancer Atlas" is based on accumulated deaths over 20 years, and does not distinguish between counties with increasing/decreasing deaths or death rates.) Identify counties with increasing trend rates believed especially significant, compare with ambient/emissions data, and discuss possible causes. Emphasis is to be placed on a method for identifying areas with increasing cancer, using ongoing local-state-Federal information-acquisition procedures.

2. Utilize Medicare Morbidity Data to Measure the Human Health Effects of Air Pollution Episodes

Using the methodology developed for analysis of the Pittsburgh air pollution episode, determine the relationships between (a) other episodes and/or selected significant temporary changes in the environment and (b) increases in hospitalizations for selected diseases. The Pittsburgh methodology is to be expanded to include emissions and ambient monitoring where data permit. Where feasible, include daily mortality by cause.

3. Determine the Existence of a Relationship Between Improving Air Quality and Improving Environmental Health

Select approximately ten counties with air quality that has improved significantly over the last decade, as measured by one or more pollutants to be selected in consultation with EPA. Select an equal number of counties, matched insofar as feasible, in all respects except for improved air quality. Select an equal number of counties on the basis of deteriorating air quality, as measured by specific pollutants. Determine hospitalizations and rates of environmentally related diseases for each set of counties. Analyze and discuss findings.

4. Determine the Existence of a Relationship Between Deteriorating Air Quality and Environmental Health

This application focuses on a relationship hypothesized as opposite to the preceding investigation. The work steps are essentially identical in concept.

5. Relationships Between Medicare Hospitalizations, Mortality and Specific Pollutants (TSP, SO_2, Respirable Suspended Particulates (RSP), etc.) for Selected Areas

Compare/contrast selected Medicare hospitalization diagnoses by county, with emissions/monitoring data for the same areas. Include provision for coordination with National Air Pollution Index monitoring sources/information. Investigate possibilities of determining threshold existence by comparisons of areas of poor air quality (exceeding standards for specific pollutants) and areas of good air quality.

a. Analyze the extensive pollutant monitoring information obtained in the multi-million dollar St. Louis project with respiratory disease hospitalizations for each county covered by the St. Louis area RAMS network. Perform a time series analysis using these data and the fine and coarse particulate fractions for elements of interest (sulfur, lead, total mass, etc.). These size fractions are readily available from the automatic dichotomous sampler data with which SSI is already working. Other readily available factors which can be taken into consideration include hourly measurements of temperature, wind speed and direction, precipitation, etc., from the RAPS data base.

b. Explore relationships between human health thresholds, and one or more pollutants. In consultation with the EPA Project Officer, prepare supplementary tabulations for unusual health status occurrences to determine the existence of a minimum level of a particular pollutant or minimum mixes of certain pollutants. Also, for selected areas/times with high hospitalization rates, investigate further the hypothesis that increased morbidity rates precede elevated mortality rates.

(1) For one selected high steel-production county (Pittsburgh, Allegheny; Birmingham, Jefferson) chart and analyze the correspondence between daily TSP/SO_2/RSP levels and hospital admissions rates for Medicare respiratory disease admissions for the same year(s). Particular attention will be paid to selected periods to be defined in consultation with EPA as pollution episodes. Outputs: Time-series plot, Julian day, and brief analysis.

(2) For Jefferson County (Birmingham), extract daily Medicare hospital admissions for respiratory diseases, for the years 1971 through 1973. For the same time period, extract TSP, RSP, and/or other data in size specifications and sampling types to be determined in consultation with EPA. See Table III-1 for type of sampling, and particle sizes. Note data date availabilities. Outputs: Julian day time series plots, and brief analyses comparing air quality with Medicare hospitalizations.

(3) Compare/contrast Birmingham respiratory disease admissions with other selected areas (not necessarily characterized by steel production concentrations) for which ambient sampling data are available. See Table III-1 for types of sampling, pollutant, areas and time periods for which the data sampling was conducted. Detailed specifications to be developed in consultation with EPA project officer. Outputs: Time series plots, tabular summaries and statistical and interpretative analyses.

6. Relationships Between Medicare Hospitalizations, Mortality, and Air Quality

a. Mortality and Medicare Morbidity Comparisons--Select disease categories believed most related to environmental conditions, and rank the counties in the United States by absolute numbers

Table III-1

TYPE OF SAMPLING, POLLUTANT, AND TIME PERIOD OF DATA FOR BIRMINGHAM, ALABAMA

Type of Sampling	Type of Analysis	Birmingham, Alabama
Hi Vol TSP 0.01 - 50 µm	TSP, NO_3^-, $SO_4^=$	10/69-9/76
Cassett 0.01 - 26 µm	RSP	6/71-9/76
HASL Cyclone AEC Curve 0% @ >10 µm 50% @ 3.5 µm 100% @ < 2 µm	RSP	6/71-9/76
Multistage Andersen Ranges 0.01 - .93 µm .93 - 1.75 µm 1.75 - 2.4 µm 2.4 - 5.5 µm 5.5 - 50 µm	RSP	4/72-2/73
CHAMP RSP Andersen 0.01 - 3.5 µm 3.5 - 26 µm	RSP, RNO_3^-, $RSO_4^=$ NRSP	12/75-9/76

of Medicare patients, total hospitalizations, and rates per resident population over 64 years of age. Compare with mortality data. The disease categories selected are recommended to be identical with one or more of the 56 categories for which SSI has already calculated age-adjusted mortality rates, by county for the approximately 10 million deaths occurring in 1968-1972.

b. Prepare an Environmental Health and Pollution Atlas of Basic Facts--

(1) Apply the multi-color county-based computer-mapping program to prepare a series of U.S. maps identifying environmental disease hospitalization prevalence (respiratory diseases, lung cancer, all neoplasms, etc.) by quintiles (or other classification), the presence of environmentally significant industries, fuel consumption by type/quantity, and other factors. Coordinate the "atlas" content with the maps and annual updating procedures of OAQPS in such manner to facilitate routine updating of environmental health status, trends, and changes.

(2) Apply the computer-mapping procedure to include ambient measures of air quality. Discuss findings, with special attention to counties with highest hospitalizations and/or highest level of pollutants.

(3) Analyze demographic characteristics of high/low counties to rationalize differences, and "map" findings.

7. Explore Relationships Between Sources and Types and Hospitalizations/ Mortality Health Indicators

Select industries of pollutant significance, or critical source types within one or more industries and compare the co-locational significance of the emissions from these sources in terms of hospitalizations and mortalities. Compare findings with areas comparable except for the absence of the industry or critical source type.

B. SUPPLEMENTARY FILES - EXTENSIONS

1. New Data

a. Obtain Medicare hospitalization files on all diseases, and more current years (the present tapes made available by HEW for 1971-1973 have only neoplasms, respiratory diseases, and diseases of the digestive system).

b. Obtain the non-confidential Social Security Administration computerized file showing the continuous work history, annually since 1957, by employer and county, of a one percent sample (about 1.5 million total workers).

2. Linkages Between Industrial/Environmental Exposures and Hospitalization/Death Diagnoses

Develop linkages, to the extent possible, between the continuous work history file and the Medicare morbidity file. It is feasible technically and at insignificant costs for both Medicare patients and the continuous work histories to be identified by the same algorithm to yield a unique, non-confidential, non-Social Security identification number. The possibilities are now being explored in HEW; complications are the novelty of the concept, and assuring compliance with bureaucratic precautions for safeguarding confidentiality.

Investigate the feasibility of utilizing the continuous work history file to improve knowledge of environmental exposures by industry and county, and relate industrial years of employment exposures to causes of hospitalization/death under Medicare.

a. A pilot investigation may build upon the project just finished by selecting known Florida counties with high in-migration Medicare hospitalization/death cases, and determining the working-life industry/locational history of individuals experiencing selected diseases such as all cancers or specific types of neoplasms.

b. Concurrently, another pilot investigation could select one or more carcinogenic hazardous industries, and for each industry select employees of one or more pre-determined age groups as of 1957. The age groups should be such as to expect Medicare hospitalizations and/or death to have occurred to a significant number of selectees since 1957, and beginning about 1970.

Conceptually, the two investigations are as follows:

(a)	(b)
For Medicare hospitalizations/deaths from selected causes of Florida Medicare residents,	For Employees of Carcinogenic Hazardous Industries of Appropriate Ages in 1957,
Determine Industrial and locational (State/County) work history.	Determine Medicare hospitalization/death causes, and migration location (State/County).

3. Death/Disability Rates by Industry/Employment - Indicators of Environmental Pollution

Using the continuous work history file, investigate death/disability rates as indicators of environmental health problems and effects among employees in industries known to be sources of environmental pollution. Include the identification of the ten (10) counties with highest and lowest death/disability rates. Relate to air quality for the counties identified. Compare rates of selected Medicare hospitalization diagnoses for the high/low counties with (a) air quality and (b) death/disability rates.

C. IMPROVEMENTS TO MEDICARE FILES IN HAND

1. Verify the Medicare hospitalizations by comparisons with morbidity files of recognized health research authorities--for example, supplement the Medicare neoplasm hospitalizations-National Cancer Institute comparison with a similar comparison of respiratory diseases with National Heart and Lung Institute files, if possible.

2. Update, and make consistent the county codes identifying Standard Metropolitan Statistical Areas to achieve compatibility with NEDS and SAROAD.

3. Identify the advantages, and difficulties, of obtaining complete consistency between Air Quality Maintenance Areas and Medicare hospitalizations occurring to residents of these areas.

4. Develop a transfer code enabling interchangeability of the Julian date of Medicare hospitalization with the more commonly used EPA system of month, day, year.

5. Sort tape by State, County, Claim Number, Admission Date to give chronological history of claimants within a state and county.

6. Sort tape by SMSA, State, County, Claim Number, Admission Date to give same chronological history as above except by SMSA.

7. Run maps similar to previous maps showing morbidity rates and other data in terms of SMSA's.

SECTION IV

DATA RESOURCES, METHODS, AND FINDINGS

This section is presented in two parts. The first part is comparatively brief, and describes the major resources consulted and the types of data assembled. The second part describes the tasks performed in this pilot study, together with observations and commentary concerning the usefulness of task products.

A. DATA RESOURCES

A major objective of this pilot study was to determine the environmental health value of certain data banks, developed for other purposes, by demonstrations with EPA's computer configuration and its associated data bases. "New" data was not to be collected. Instead, emphasis was placed on exploiting existing information sources, preferably computer-encoded.

The following description covers the major data files used in this project. Included are example printouts. Incidental sources of information are included as appropriate in the text discussing a specific demonstration or application.

1. Medicare Files

The approximately 25 million Medicare beneficiaries who are provided medical treatment under Social Security represent the largest population group for which morbidity data are collected regularly and compiled systematically nation-wide by a single program. This group includes essentially 100 percent of the population 65 years of age and older. Collectively, this group experiences about 5.5 million hospital admissions annually. They account for about half the hospital cancer cases in the United States.

System Sciences, Inc. (SSI) obtained tape reels from the Social Security Administration containing data from the 1971, 1972, and 1973 Medicare short-stay hospital sample discharge survey. Table IV-1 shows the record layout for each of the tapes. The State and County codes of field numbers 4 and 5 pertain to the place of residence of the claim beneficiary. Records for which the discharge diagnosis is "neoplasm" (ICDA 140-239), "disease of the respiratory system" (ICDA 460-519), or "disease of the digestive system" (ICDA 520-577) are included in these files; records with other codes are available from the Social Security Administration, but were not deemed necessary for purposes of the pilot study.

There are approximately 400,000 admission/discharge records for each of the three years. This represents a 20 percent sample of qualifying claims and includes multiple records for some individuals. The Social Security numbers are scrambled in accordance with a single, unique algorithm to permit a longitudinal matching of individual claim histories, by means other than use of the Social Security number. This provision enables assignment to each Medicare beneficiary of a unique alphanumeric identifier, but with the individual's privacy and identity fully protected. The Medicare sample selectee procedure is "permanent"; that is, once in the sample, the selectee's Medicare history is routinely up-dated until death through the claims/insurance/payment process. Also, the longitudinal file starts with the first Medicare claim for each individual.

2. POPATRISK Data Base

A data base, called "POPATRISK,"* was developed by SSI to provide the EPA with a user-oriented data base containing recent county-based information for all counties in the contiguous United States, including population demographics, population mobility, climatology, emissions, air quality, and age-adjusted death rates.

---

*Population at Risk to Various Air Pollution Exposures: Data Base "POPATRISK"; Contract No. 68-02-2269, for Health Effects Research Laboratory, EPA/Research Triangle Park, North Carolina. EPA-600/1-78-051, June 1978.

Table IV-1. RECORD LAYOUT FOR DATA TAPES FROM MEDICARE HOSPITAL DISCHARGE SURVEY

Field Name	Field Size & Type	Positions Beg	Positions End	Contents of Field
1. Claim number	14 alphanumeric	1	14	
2. Sex	1		15	1 - Male 2 - Female 3 - Unknown
3. Race	1		16	1 - white; 2 - negro; 3 - other; 4 - unknown
4. State	2	17	18	SSA state and county code
5. County	3	19	21	SSA state and county code
6. Admission Date	5	22	26	Julian Date YYDDD
7. Discharge Status	1		27	0 - alive; 1 - dead
8. Additional Diagnoses	1		28	1 - more than 1; 0 - none
9. Discharge Diagnoses	4	29	32	ICDA (Codes 140-239 and 460-577)
10. SMSA - beneficiary	3	33	35	Blank - not in SMSA; 999 - unknown, or actual 3-digit SMSA code
11. Age	3	36	38	
12. Spaces	4	39	42	

"POPATRISK" contains approximately 27.5 million characters and is in SYSTEM 2000, Version 2.80 format, facilitating access with minimal user computer training. Population demographics are as of the 1970 Census; population mobility is described spanning the years 1965 to 1970 for six sex-race categories in seven age groupings for both "in" and "out" migrants; climatology information contains county summaries of temperature, precipitation and hours of sunshine; county point and area source emission estimates are provided for five criteria pollutants--TSP, SO_2, NO_2, CO, and ozone--based on the NEDS-USER file; air quality information is based on 1974 data contained in SAROAD; age-adjusted death rates were computed for the combined years 1969, 1970, and 1971 for four sex-race categories in 50 groupings of ICDA categories (eighth revision).

3. "Industrial Correlates" Project Statistics

Under contract with the EPA Office of Toxic Substances, SSI systematically related mortality rates of specified disease categories to certain aspects of the environment with the primary focus being upon the relationship between community proximity to industrial effluents/emissions and the various mortality rates. One of the products of that research effort was a ranking of the white male and white female age-adjusted death rates (per 100,000; ages 34-75) in all the counties of the United States. A "hard copy" listing of the top 500 counties in each disease category was also prepared. An example ranking of the top U.S. counties by white male mortality rates for disease category 04 (all neoplasms--ICDA 140-239) is displayed in Table IV-2.

4. County Business Patterns

A primary data resource for documenting industrial development in the study area was County Business Patterns, published by the U.S. Department of Commerce and based on Social Security Administration reports of first-quarter earnings. Industry sectors are classified by 2-, 3-, and 4-digit Standard Industrial Classification codes. A sample page of data for Pinellas County, Florida, 1975, is shown in Table IV-3.

Table IV-2

COUNTIES WITH HIGHEST WHITE MALE MORTALITY RATES FOR ALL NEOPLASMS (ICDA 140-239)
(Ages 34-75, Rates per 100,000)

RANK	RATE	ST-COU	NAME
001.	1310	55-078	MENOMONEE, WIS
002.	852	06-003	ALPINE, CAL
003.	790	08-079	MINERAL, COL
004.	770	30-069	PETROLEUM, MON
005.	767	38-007	BILLINGS, N D
006.	721	12-129	WAKULLA, FLA
007.	697	25-019	NANTUCKET, MAS
008.	654	13-183	LONG, GA
009.	653	13-049	CHARLTON, GA
010.	652	12-035	FLAGLER, FLA
011.	622	16-025	CAMAS, IDA
012.	622	26-079	KALKASKA, MIC
013.	610	38-085	SIOUX, N D
014.	608	13-061	CLAY, GA
015.	588	13-253	SEMINOLE, GA
016.	574	48-311	MC MULLEN, TEX
017.	564	46-075	JONES, S D
018.	561	32-027	PERSHING, NEV
019.	555	12-041	GILCHRIST, FLA
020.	549	51-043	CLARKE, VA
021.	548	01-113	RUSSELL, ALA
022.	546	22-087	ST. BERNARD, LA
023.	534	54-065	MORGAN, W V
024.	533	45-065	MCCORMICK, S C
025.	530	28-161	YALOBUSHA, MIS
026.	523	51-127	NEW KENT, VA
027.	520	51-137	ORANGE, VA
028.	520	08-053	HINSDALE, COL
029.	517	13-029	BRYAN, GA
030.	509	51-181	SURRY, VA
031.	509	51-131	NORTHAMPTON, VA
032.	507	24-037	ST. MARYS, MD
033.	506	28-103	NOXUBEE, MIS
034.	499	13-007	BAKER, GA
035.	496	17-003	ALEXANDER, ILL
036.	495	30-037	GOLDEN VALLEY, MON
037.	493	24-007	BALTIMORE CITY, MD
038.	492	17-153	PULASKI, ILL
039.	490	32-015	LANDER, NEV
040.	486	20-021	CHEROKEE, KAN
041.	485	13-237	PUTNAM, GA
042.	481	48-313	MADISON, TEX
043.	480	51-053	DINWIDDIE, VA
044.	479	28-033	DE SOTO, MIS
045.	478	40-115	OTTAWA, OKL
046.	478	22-101	ST. MARY, LA
047.	476	13-173	LANIER, GA
048.	476	39-121	NOBLE, OH
049.	476	28-011	BOLIVAR, MIS
050.	475	21-193	PERRY, KY

Reproduced from best available copy.

Table IV-3. SAMPLE PAGE

Table 2. Counties—Employees, Payroll, and Establishments, by Industry: 1975—Continued

(Excludes government employees, railroad employees, self-employed persons, etc.—see "General Explanation" for definitions and statement on reliability of data. Size class 1 to 4 includes establishments having payroll but no employees during mid-March pay period. "D" denotes figures withheld to avoid disclosure of operations of individual establishments; the other alphabetics indicate employment-size class—see footnote.)

SIC code	Industry	Number of employees for week including March 12	Payroll ($1000) First quarter	Payroll ($1000) Annual	Number of establishments, by employment-size class: Total	1 to 4	5 to 9	10 to 19	20 to 49	50 to 99	100 to 249	250 to 499	500 to 999	1000 or more
	PINELLAS—Continued													
344	Fabricated structural metal products	1 090	2 720	9 506	35	9	5	10	7	1	2	1	-	-
3441	Fabricated structural metal	152	424	1 913	6	1	1	2	1	1	-	-	-	-
3442	Metal doors, sash, and trim	59	102	385	3	-	-	1	2	-	-	-	-	-
3443	Fabricated plate work (boiler shops)	(C)	(D)	(D)	2	1	-	-	-	-	1	-	-	-
3444	Sheet metal work	548	1 048	4 225	14	3	2	5	3	-	-	1	-	-
3446	Architectural metal work	54	117	438	7	4	1	1	1	-	-	-	-	-
3448	Prefabricated metal buildings	(C)	(D)	(D)	2	-	-	1	-	-	1	-	-	-
345	Screw machine products, bolts, etc	125	367	1 313	7	1	1	2	3	-	-	-	-	-
3451	Screw machine products	(C)	(D)	(D)	6	1	-	2	3	-	-	-	-	-
347	Metal services, nec	159	335	1 213	11	3	4	2	2	-	-	-	-	-
3471	Plating and polishing	136	291	1 054	7	1	2	2	2	-	-	-	-	-
349	Misc. fabricated metal products	203	503	1 717	10	5	3	-	1	-	1	-	-	-
3499	Fabricated metal products, nec	(C)	(D)	(D)	9	4	3	-	1	-	1	-	-	-
35	Machinery, except electrical	2 005	5 431	21 730	80	28	19	17	7	7	3	1	-	-
353	Construction and related machinery	(G)	(D)	(D)	4	1	-	-	2	1	-	-	-	-
3535	Conveyors and conveying equipment	(B)	(D)	(D)	1	-	-	-	-	1	-	-	-	-
354	Metalworking machinery	544	1 720	6 974	17	3	4	5	1	3	1	-	-	-
3542	Machine tools, metal forming types	(C)	(D)	(D)	1	-	-	-	-	-	1	-	-	-
3544	Special dies, tools, jigs & fixtures	310	843	3 826	10	1	2	4	-	3	-	-	-	-
3545	Machine tool accessories	(B)	(D)	(D)	5	2	1	1	1	-	-	-	-	-
355	Special industry machinery	272	866	2 803	8	1	2	2	2	-	1	-	-	-
3551	Food products machinery	(C)	(D)	(D)	3	-	1	-	1	-	1	-	-	-
356	General industrial machinery	321	1 131	4 994	8	2	2	1	-	2	-	1	-	-
3564	Blowers and fans	(B)	(D)	(D)	2	1	-	-	-	1	-	-	-	-
3569	General industrial machinery, nec	(C)	(D)	(D)	4	-	1	1	-	1	-	1	-	-
358	Refrigeration and service machinery	(B)	(D)	(D)	5	4	-	-	-	1	-	-	-	-
3589	Service industry machinery, nec	(B)	(D)	(D)	1	-	-	-	-	1	-	-	-	-
359	Misc. machinery, except electrical	448	1 161	4 754	33	12	10	8	2	-	1	-	-	-
3599	Machinery, except electrical, nec	448	1 161	4 754	33	12	10	8	2	-	1	-	-	-
36	Electric and electronic equipment	8 187	25 950	101 635	34	10	6	3	7	1	-	3	1	3
361	Electric distributing equipment	(F)	(D)	(D)	2	1	-	-	-	-	-	-	1	-
3612	Transformers	(F)	(D)	(D)	2	1	-	-	-	-	-	-	1	-
366	Communication equipment	(I)	(D)	(D)	7	-	1	1	1	-	-	2	-	2
3662	Radio and tv communication equipment	(I)	(D)	(D)	6	-	1	-	1	-	-	2	-	2
367	Electronic components and accessories	(G)	(D)	(D)	12	2	1	1	5	1	-	1	-	1
3676	Electronic resistors	(E)	(D)	(D)	2	-	-	-	-	1	-	1	-	-
3679	Electronic components, nec	(G)	(D)	(D)	6	-	1	1	3	-	-	-	-	1
37	Transportation equipment	1 415	3 333	12 836	41	18	6	5	4	4	3	1	-	-
372	Aircraft and parts	(E)	(D)	(D)	3	-	1	-	-	1	1	-	-	-
3728	Aircraft equipment, nec	(E)	(D)	(D)	2	-	-	-	-	1	1	-	-	-
373	Ship and boat building and repairing	1 033	2 309	9 319	29	15	4	1	3	3	2	1	-	-
3731	Ship building and repairing	(E)	(D)	(D)	3	1	-	-	-	1	1	-	-	-
3732	Boat building and repairing	(F)	(D)	(D)	26	14	4	1	3	2	1	1	-	-
379	Miscellaneous transportation equipment	(B)	(D)	(D)	6	1	-	4	1	-	-	-	-	-
38	Instruments and related products	817	1 984	4 957	22	10	5	2	-	2	3	-	-	-
384	Medical instruments and supplies	479	1 550	3 584	14	7	3	1	-	1	2	-	-	-
3841	Surgical and medical instruments	(E)	(D)	(D)	4	2	-	-	-	-	2	-	-	-
3842	Surgical appliances and supplies	(C)	(D)	(D)	8	4	2	1	-	1	-	-	-	-
385	Ophthalmic goods	(B)	(D)	(D)	1	-	-	-	-	1	-	-	-	-
39	Miscellaneous manufacturing industries	407	883	3 153	46	24	12	6	3	-	1	-	-	-
394	Toys and sporting goods	106	162	738	12	5	2	4	1	-	-	-	-	-
3949	Sporting and athletic goods, nec	(B)	(D)	(D)	10	4	2	3	1	-	-	-	-	-
395	Pens, pencils, office and art supplies	(C)	(D)	(D)	4	1	2	-	-	-	1	-	-	-
3951	Pens and mechanical pencils	(C)	(D)	(D)	1	-	-	-	-	-	1	-	-	-
399	Miscellaneous manufactures	146	350	1 470	24	14	7	2	1	-	-	-	-	-
3993	Signs and advertising displays	129	264	1 122	18	10	5	2	1	-	-	-	-	-
—	Administrative and auxiliary	531	1 442	6 186	7	-	1	1	1	1	3	-	-	-
—	Transportation and other public utilities	8 728	22 249	93 011	228	108	45	39	16	14	2	2	-	2
41	Local and interurban passenger transit	469	1 200	4 717	35	8	12	8	4	3	-	-	-	-
411	Local and suburban transportation	298	450	1 773	16	1	7	5	2	1	-	-	-	-
4111	Local and suburban transit	(C)	(D)	(D)	2	-	-	-	1	1	-	-	-	-
4119	Local passenger transportation, nec	(C)	(D)	(D)	14	1	7	5	1	-	-	-	-	-
412	Taxicabs	(C)	(D)	(D)	17	7	5	3	1	1	-	-	-	-
413	Intercity highway transportation	(E)	(D)	(D)	2	-	-	-	1	1	-	-	-	-
42	Trucking and warehousing	1 061	1 935	8 286	73	35	14	13	7	3	1	-	-	-
421	Trucking, local and long distance	(G)	(D)	(D)	71	34	14	12	7	3	1	-	-	-
44	Water transportation	131	231	1 033	24	16	4	4	-	-	-	-	-	-
446	Water transportation services	(C)	(D)	(D)	20	14	2	4	-	-	-	-	-	-
4469	Water transportation services, nec	(C)	(D)	(D)	20	14	2	4	-	-	-	-	-	-
47	Transportation services	251	415	1 802	40	27	8	3	1	1	-	-	-	-
472	Arrangement of transportation	(C)	(D)	(D)	38	26	8	3	-	1	-	-	-	-
4722	Passenger transportation arrangement	(C)	(D)	(D)	38	26	8	3	-	1	-	-	-	-
48	Communication	(H)	(D)	(D)	24	9	4	5	3	2	-	-	-	1
483	Radio and television broadcasting	284	617	2 649	11	-	3	4	3	1	-	-	-	-
489	Communication services, nec	(B)	(D)	(D)	4	3	-	-	-	1	-	-	-	-
49	Electric, gas, and sanitary services	(H)	(D)	(D)	25	9	2	4	1	5	1	2	-	1
491	Electric services	(G)	(D)	(D)	8	-	-	1	1	4	1	-	-	1

A 0-19. B 20-99. C 100-249. E 250-499. F 500-999. G 1 000-2 499. H 2 500-4 999. I 5 000-9 999. J 10 000-24 999. K 25 000-49 999. L 50 000-99 999. M 100 000 or more

5. NEDS Emissions Inventory

One project task required identification of counties with steel-production facilities and determination of the fraction of point source emissions attributable to these sources. Retrievals were made from EPA's National Emissions Data System for information about individual SIC 3312 point-source emissions of Total Suspended Particulate and Sulfur Oxides. An illustration of the data obtained is shown in Table IV-4.

B. METHODS AND FINDINGS

This second and major part of the Data Resources, Methods and Findings section describes the tasks performed in this pilot study and principal results or findings, as well as appropriate observations and comparisons useful in assessing the utility of the task products. A discussion of initial data processing is presented, as well as the general characteristics of the sample population, and a description of the methodology for choosing study areas and documenting their industrial development. The section concludes with descriptions of three investigations undertaken in this project; (1) a comparison of the availability and uses of morbidity and mortality data, (2) evaluation of the significance of Medicare claimant migration, and (3) a demonstration of morbidity/mortality analyses in steel industry counties.

The latter demonstration features computer-prepared multi-color United States maps displaying, by county, rates of Medicare hospitalizations for respiratory disease together with the relative importance of the steel industry as a source of TSP emissions. This inexpensive technique not only facilitates analyses, but is extremely effective in communicating the complexities of environmental health interactions.

1. Initial Processing of the Medicare Data Files

The tape reels containing data from the 1971, 1972, and 1973 Medicare short-stay hospital sample discharge survey were received from the Social Security Administration by the end of the first month of the contract performance period. There was a single tape for each of the three years. The program documentation accompanying the files indicated 373,877, 408,204,

Table IV-4

SAMPLE POINT SOURCE LISTINGS FROM NATIONAL EMISSIONS DATA SYSTEM
FOR PARTICULATE AND SULFUR OXIDES EMISSIONS (IN SHORT TONS, 1977)

```
                                                                                 -PART-- --S O X--

-0039: U S STEEL-GARY 46401 PART 2-  -----------OWNERSHIP: PRIVATE-----   ---2---------82---
 15: INDIANA          2360: LAKE CO                YEAR OF RECORD:  1977
-067: METROPOLITAN CHICAGO
 000: NO CONTROL EQUIPMENT                     EFF =  0.0 %  POINT: 34
-ESTIMATE BY (2) MATERIALS BALANCE    SIC = 3312  SCC = 1-02-007-08

-0039: U S STEEL-GARY 46401 PART 2-  -----------OWNERSHIP: PRIVATE-----   ---1---------42---
 15: INDIANA          2360: LAKE CO                YEAR OF RECORD:  1977
 067: METROPOLITAN CHICAGO
 000: NO CONTROL EQUIPMENT                     EFF =  0.0 %  POINT: 35
 ESTIMATE BY (2) MATERIALS BALANCE    SIC = 3312  SCC = 1-02-007-08

-0039: U S STEEL-GARY 46401 PART 2-  -----------OWNERSHIP: PRIVATE-----   ---1---------62---
 15: INDIANA          2360: LAKE CO                YEAR OF RECORD:  1977
-067: METROPOLITAN CHICAGO
 000: NO CONTROL EQUIPMENT                     EFF =  0.0 %  POINT: 36
 ESTIMATE BY (2) MATERIALS BALANCE    SIC = 3312  SCC = 1-02-007-08

-0039: U S STEEL-GARY 46401 PART 2-  -----------OWNERSHIP: PRIVATE-----   ---1---------54---
 15: INDIANA          2360: LAKE CO                YEAR OF RECORD:  1977
-067: METROPOLITAN CHICAGO
 000: NO CONTROL EQUIPMENT                     EFF =  0.0 %  POINT: 37
-ESTIMATE BY (2) MATERIALS BALANCE    SIC = 3312  SCC = 1-02-007-08

-0039: U S STEEL-GARY 46401 PART 2-  -----------OWNERSHIP: PRIVATE-----   ---2---------101--
 15: INDIANA          2360: LAKE CO                YEAR OF RECORD:  1977
-067: METROPOLITAN CHICAGO
 000: NO CONTROL EQUIPMENT                     EFF =  0.0 %  POINT: 38
 ESTIMATE BY (2) MATERIALS BALANCE    SIC = 3312  SCC = 1-02-007-08
```

and 414,434 records for the respective years. These numbers were independently verified as identical with the 20 percent sample of records discussed in this report.

A computer account was established at the EPA National Computation Center in Research Triangle Park, and the Medicare data tapes were registered there. The data were copied onto system tapes, and the original reels were released to SSI. An initial reading of the tapes assured the readability and appropriate formatting of the various data-field items.

To allow statistical computations of sample characteristics on a per-patient, as well as a per-record, basis, and to develop a chronology for individual claimants regarding diagnosis, discharge status, beneficiary residence, etc., the three years of Medicare records were merged and sorted by claim number. A sample of the longitudinally merged and sorted records, grouped by claim numbers, is shown in Table IV-5. Each line in the table is one record, representing a single admission and discharge. The claim number consists of a twelve-digit number obtained by scrambling the Social Security number under which a claim was filed, followed by a one-or-two-place suffix identifying the beneficiary involved. Beneficiary identification codes are listed in Tables IV-6 and IV-7. Thus a unique claim number represents a specific individual. The records in Table IV-5 are further ordered by date of admission to provide a chronological history for each individual, for the years 1971, 1972, and 1973.

2. Sample Characteristics

Following creation of the longitudinal file, counts were taken of the items shown in Tables IV-8, IV-9, and IV-10. Table IV-8 indicates that the consolidated data file contains about 1.2 million admission/discharge records (representing a sum of records for the three years) filed under 815,000 unique claim numbers. Although there may be instances of individuals filing for benefits under more than one Social Security number, one would expect the proportion of such cases to be negligible and that 4,075,000 (five times 815,000) is a good approximation to the number of individuals in the United States treated as in-patients under Medicare provisions. There were 1,196,000

TABLE IV-5. SAMPLE DISPLAY OF MEDICARE DATA ASSEMBLED LONGITUDINALLY TO SHOW HOSPITALIZATION CAUSES, 1971-1973 1/

Claim Number	Sex[a]	Race[b]	State	County	Julian Date Admission	Discharge Status[c]	Additional Diagnosis[d]	Discharge Diagnosis	SMSA	Age
000072502693J2	1	1	05	150	71123	0	1	493X		073
000072502693J2	1	1	05	150	71156	0	0	493X		073
000072502693J2	1	1	05	150	71184	0	0	493X		073
000072502693J2	1	1	05	150	71362	0	0	493X		074
000072502693J2	1	1	05	150	72041	0	1	493X		074
000072502693J2	1	1	05	150	72364	0	1	5621		075
000072502693J2	1	1	05	150	73183	1	1	491X		075
000074178949A	1	1	05	310	73305	0	1	1621	526	065
000074178949A	1	1	05	310	73308	0	1	1621	526	065
000074215383A	1	1	05	480	73157	0	0	5192	526	065
000074215383A	1	1	05	480	73157	0	1	5192	526	065
000074215383A	1	1	05	480	73190	0	1	5192	526	065
000074308375A	2	1	05	480	71229	0	0	174X	526	078
000074597093D	2	1	05	530	71189	0	1	1972	529	077
000074597093D	2	1	05	530	71229	1	1	1621	529	077

[a] Sex: 1=male, 2=female, 3=unknown

[b] Race: 1=white, 2=black, 3=other, 4=unknown

[c] Discharge Status: 0=alive, 1=dead

[d] Additional Diagnosis: 0=none, 1=more than one

1/ To illustrate the wealth of data and linkages, the first individual is eligible as a primary SSA beneficiary ("J" in claim number), a white male, with 7 admissions, 4 in 1971, 2 in 1972, and a third in 1973 which ended in death. During this period, he did not change his county of residence, which is not a metropolitan area. His first 5 admissions were for asthma (493X); his sixth admission was diverticulitis of the colon (562.1); and approximately six months later his seventh admission terminated in death from chronic bronchitis (491X). Additional diagnoses of an unspecified nature were made for his first admission, and his last three admissions. The last example shown is eligible as a widow ("D" in claim number), white, with two admissions in 1971, and a resident of a metropolitan area. Her first admission was classed as secondary malignant neoplasm of respiratory system, pleura (197.2); and her second admission, 40 days later, terminated in death, cause 162.1, malignant neoplasm of bronchus and lung.

This display is by claim number, within state. The data may be re-grouped by county, by diagnosis, by admission date, or other categories to facilitate analyses.

TABLE IV-6

BENEFICIARY IDENTIFICATION CODES

SOCIAL SECURITY ADMINISTRATION TYPE OF CLAIM

A	Wage Earner (Old-age or Disability)
B	Aged Wife
B3	Second Claimant
B8	Third Claimant
B1	Husband
B4	Second Claimant
B2	Young Wife (entitled child in her care)
B5	Second Claimant
B7	Third Claimant
B6	Divorced Wife
B9	Second Claimant
C1	Children (youngest child)
C2	(next to youngest child, etc.) (C10 would appear as CA, C11 as CB, C12 as CC, etc.)
D	Aged Widow
D2	Second Claimant
D8	Third Claimant
D4	Remarried Widow (60 years of age or older)
D1	Widower
D3	Second Claimant
D5	Remarried Widower (62 years of age or older)
D6	Surviving Divorced Wife
D7	Second Claimant
E	Mother (Widow)
E2	Second Claimant
E7	Third Claimant
E1	Surviving Divorced Mother
E3	Second Claimant
F1	Father
F2	Mother
F3	Stepfather
F4	Stepmother
F5	Adopting Father
F6	Adopting Mother
F7	Second Alleged Father
F8	Second Alleged Mother
J1	Primary (less than 3QC's and entitled to HIB)
J2	(3 or more QC's and entitled to HIB)
J3	(less than 3 QC's and not entitled to HIB)
J4	(3 or more QC's and not entitled to HIB)
K1	Wife's (less than 3QC's and entitled to HIB)
K2	(3 or more QC's and entitled to HIB)
K3	(less than 3QC's and not entitled to HIB)
K4	(3 or more QC's and not entitled to HIB)
M	Uninsured (Not entitled to HIB, but qualified for SMIB)
M1	Uninsured (Qualifies for HIB but requests only SMIB)
T	HIB Entitlement (deemed insured)

TABLE IV-7

BENEFICIARY IDENTIFICATION CODE

RAILROAD BOARD EMPLOYEE TYPE OF CLAIM

10	Railroad employee annuitant
11	Survivor joint annuitant (election to reduced annuity to guarantee widow payment)
13, 43, or 83 *	Widow with child in her care or child alone of a deceased railroad employee
14 or 84 *	Spouse of a railroad employee
15 or 85 *	Parent of a railroad employee
16, 46, or 86 *	Widow or widower of a railroad employee

*1 in subscript denotes RR annuitant involved
8 in subscript denotes RR pensioner involved

TABLE IV-8

PATIENT NON-MEDICAL CHARACTERISTICS
FROM CONSOLIDATED MEDICARE RECORDS FOR 1971-1973

Item	Count	%
Admission/Discharge records	1,196,515	--
Unique claim numbers	815,000	100.00
White males	359,017	44.05
White females	386,674	47.44
Non-white males	25,843	3.17
Non-white females	24,589	3.02
Unknown race and/or sex	18,877	2.32
Discharged Alive	710,323	87.16 (100.00)
1 admission	527,231	(74.22)
2 admissions	121,322	(17.08)
3 admissions	36,965	(5.20)
4 admissions	13,388	(1.88)
5 admissions	5,591	(0.79)
6 or more admissions	5,826	(0.82)
Discharged Dead	104,677	12.84 (100.00)
1 admission	58,612	(55.99)
2 admissions	25,385	(24.25)
3 admissions	11,352	(10.84)
4 admissions	4,986	(4.76)
5 admissions	2,176	(2.08)
6 or more admissions	2,166	(2.07)
Additional diagnosis (coded 1 - yes)	737,613	61.65
Age 65 or younger in 1971	102,400	12.56 (100.00)
Males	52,757	(51.52)
Females	49,643	(48.48)
Railroad Retirement beneficiaries	26,596	3.26

TABLE IV-9

DISEASE PREVALENCE FROM CONSOLIDATED MEDICARE RECORDS FOR 1971-1973

Item	Count	%
Discharge Diagnosis (Neoplasms)		*
ICDA 140-145, buccal cavity	4,439	0.54
ICDA 146-149, pharynx	1,479	0.18
ICDA 151, stomach	7,290	0.89
ICDA 153, large intestine (exc. rectum)	25,843	3.17
ICDA 154, rectum/rectosigmoid junction	10,334	1.27
ICDA 157, pancreas	6,177	0.76
ICDA 150, 152, 155, 156, 158, 159, other digestive	6,352	0.78
ICDA 162, trachea/bronchus/lung	23,409	2.87
ICDA 160, 161, 163, other respiratory	4,133	0.51
ICDA 170-174, bone/connective tissue/ skin/breast	38,318	4.70
ICDA 185, prostate	24,652	3.02
ICDA 188, bladder	13,085	1.61
ICDA 180-184, 186, 187, 189, other genitourinary	18,145	2.23
ICDA 191, brain	960	0.12
ICDA 190, 192-199, other and unspecified	25,202	3.09
ICDA 204-207, leukemia	6,450	0.79
ICDA 200-203, 208, 209, other lymphatic and hematopoietic tissue	10,905	1.33
ICDA 210-228, benign	29,668	3.64
ICDA 230-239, unspecified nature	19,634	2.41
Discharge Diagnosis (Respiratory Diseases)		
ICDA 460-474, acute infection/influenza	57,615	7.07
ICDA 480-486, pneumonia	100,347	12.31
ICDA 490-493, bronchitis/emphysema/asthma	62,465	7.66
ICDA 500-519, other respiratory	66,584	8.17
Discharge Diagnosis (Digestive Diseases)		
ICDA 530-537, esophagus/stomach/duodenum	85,119	10.44
ICDA 570-577, liver/gallbladder/pancreas	95,468	11.71
ICDA 520-529, 540-569, other	233,110	28.60
	977,183	119.90*

---

* A unique claim number will be counted no more than once within a disease classification category, but a claim history (with more than one admission/discharge) may contain more than one category of discharge diagnosis.

TABLE IV-10

DISEASE CATEGORY COMBINATIONS FROM CONSOLIDATED MEDICARE RECORDS FOR 1971-1973

Item	Count	%
Unique claim numbers	815,000	100.00
Unique claim numbers for which complete admission/discharge hisotry shows:		
Neoplasm diagnoses only	200,902	24.65
Respiratory disgnoses only	208,527	25.59
Digestive diagnoses only	330,022	40.49
Neoplasm and respiratory diagnoses	15,441	1.89
Respiratory and digestive diagnoses	30,486	3.74
Neoplasm and digestive diagnoses	26,746	3.28
Neoplasm, respiratory, and digestive diagnoses	2,876	0.35

admissions, about 1.5 admissions per claimant. This compares with an average of 1.7 admissions per patient during a two-year follow-up period found by the Third National Cancer Survey.

More than 91 percent of the claimants were white, with white males accounting for about 44 percent of the sample (or 359,017 individuals). Close to 13 percent of the sampled individuals admitted to hospitals for treatment were discharged dead; there were approximately 28 percent more admissions for those eventually discharged dead during the sample years than for those discharged alive following each recorded hospital admission.

Nearly 62 percent of the records showed a diagnosis in addition to the ICDA-coded discharge diagnosis. As the presence of other condition(s) is indicated, but not its disease code, this data field cannot now contribute to environmental health effects analyses until coded. It constitutes an additional epidemiological resource in an area where comprehensive and uniformly defined statistics are generally inadequate for environmental analyses.

There were 102,400 claimants (12.56 percent of the sample) who were 65 years of age in 1971 or later and were probably making their first claims under Medicare. A field on the unabstracted files maintained by the Social Security Administration is labeled "Formerly Disabled" and could provide some further insights on this point.

The claim number also identifies railroad employees and the spouses or surviving beneficiaries covered by health insurance provided under the Railroad Retirement Act (RRA). Although this industry is not of primary research interest to this project, it was thought that the suffixes to claim numbers in the Medicare data identifying RRA claims offer an excellent opportunity to compare the morbidity/mortality differentials of working husbands and their wives on a nationwide basis for a single industry.

Two types of analysis were performed using the four-digit ICDA-coded discharge diagnosis. In one procedure the neoplasms, respiratory and digestive disorders were partitioned into 19, four and three categories, respectively, and the numbers of claimants whose hospitalization showed at least one admission for a category were registered. This provided a disease prevalence indicator, and the sample distribution is presented in Table IV-9. The cancer diagnosis code groupings were chosen to incorporate cancer types with high incidence and relatively low five-year survival rates. The lower-survival-rate (or high death rate) consideration was of importance since only three years of Medicare morbidity data were available with which to facilitate morbidity/mortality comparisons.

The second procedure characterizes the complete hospitalization history associated with each claim number by whether neoplasm, respiratory, or digestive disorders have occurred exclusively or in combination with another type of disorder. For example, a claimant with a history of admissions for cancer may be classified as neoplasm diagnoses only, neoplams and respiratory diagnoses, neoplasm and digestive diagnoses, or neoplasm, respiratory, and digestive diagnoses. Results are shown in Table IV-10.

Of the 815,000 Medicare claimants in the hospitalization sample, 245,965 had an admission for which the discharge diagnosis was "neoplasm" (ICDA 140-239); 257,330 had an admission for respiratory disease; and 390,130 were admitted for a digestive disease. Since this Medicare file includes records for only three stated disease groupings, it is possible that admission/discharge histories with a neoplasm and a non-abstracted disease classification, such as cardiovascular disorder, are counted in the "neoplasm diagnoses only" category. It is interesting to note that 90.73 percent of the individuals in the sample did not have diagnoses in more than one of the abstracted categories (neoplasm, respiratory, and digestive).

3. Choosing the Study Areas for Selected Morbidity/Mortality Comparisons

To select counties of particular interest in this study of correspondences between area morbidity and mortality statistics, retrievals from the POPATRISK data base were made, and other data resources were examined. Information concerning population distribution, migration, air quality, and industrial employment were among the factors considered. For example, POPATRISK contains Census migration survey data covering the years 1965-1970, from which counties with stable populations and those with significant changes in the number of residents over that time period can be determined. It was considered particularly important to differentiate between "retirement" counties characterized by large fractions of older persons who had lived/worked elsewhere during their earlier years, and "non-retirement" counties with a younger and more stable population.

In consultation with EPA, attention was focused on U.S. counties having 1970 Census population greater than 10,000. Within this population specification, 31 counties with low migration (both in- and out-) for white males of all ages, and 45 counties with comparably low in- and out-migration for white males over the age of 44 were identified. There are 15 counties which appear on both lists. The maps in Figs. IV-1 and IV-2 show the locations of the counties meeting these criteria.

Tables IV-11 and IV-12 indicate the counties with high migration, either in- or out-, among white males over 44, and with large differences between rates of in-migration and out-migration. More specifically, Table IV-11 lists counties with ratio of in/out migration greater than 3.0. For the same population category, Table IV-12 shows the counties with an out/in migration ratio greater than 2.0.

Because the older age groups are of primary interest, and because changes in morbidity and mortality statistics may be more easily attributable to population increases due to in-migration than to losses of out-migrants, only the 75 counties of Fig. IV-2 and Table IV-11 were considered further. For those 75 counties, absolute numbers of white males over the age of 64 and the percentage of that group which changed county of residence between 1965 and 1970 were obtained from POPATRISK.

The POPATRISK file also provided emissions and monitoring measures for TSP and SO_2. The emissions values are NEDS estimates expressed in tons per year; the air quality values are the arithmetic means, for each county, of the annual geometric means of data from individual population-based monitoring sites. In cases where monitoring data were not available, a regression equation using NEDS data for TSP and oxides of nitrogen (NO_x) was used to estimate a county air quality index with regard to TSP.

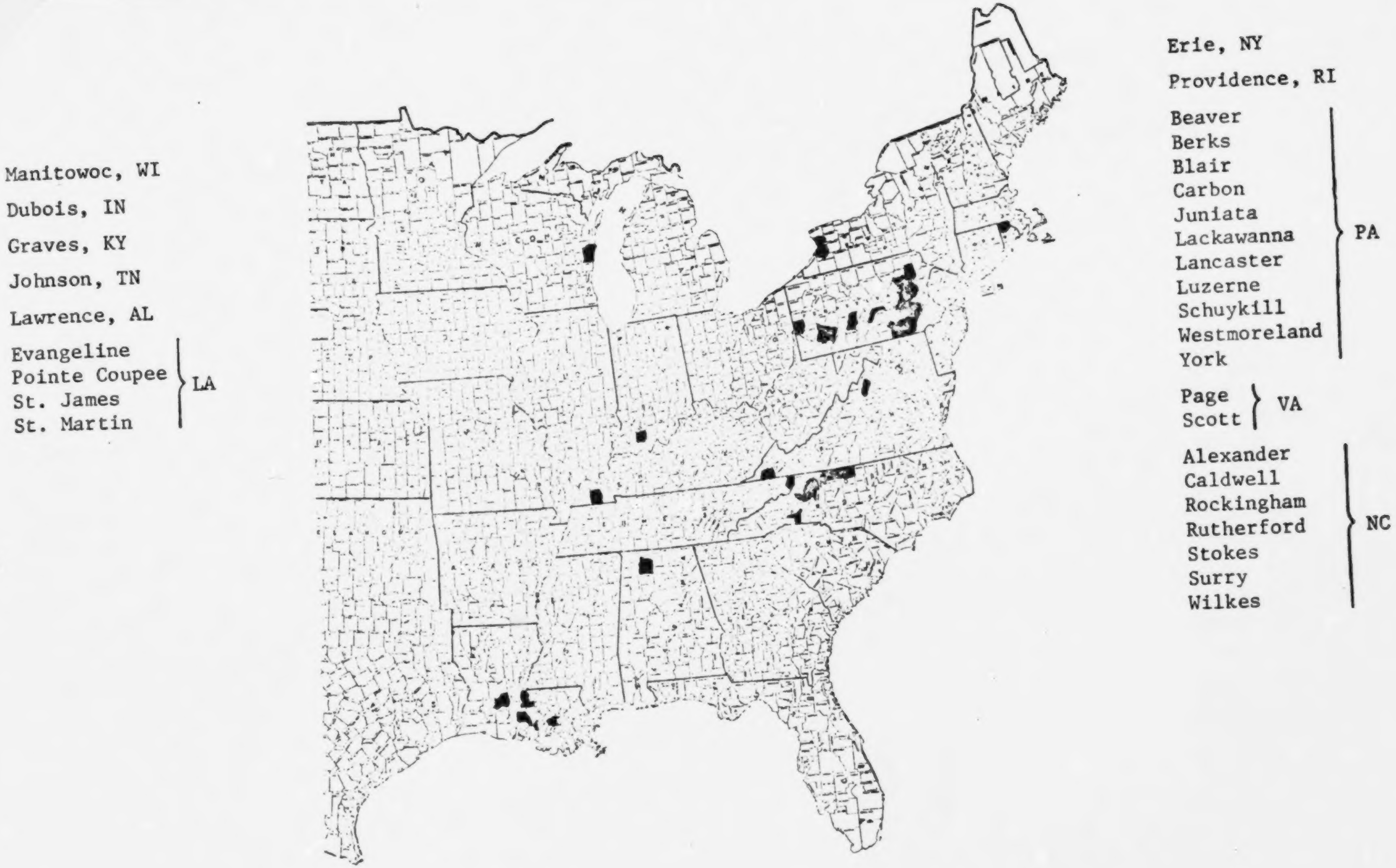

Manitowoc, WI
Dubois, IN
Graves, KY
Johnson, TN
Lawrence, AL
Evangeline
Pointe Coupee
St. James
St. Martin
LA
Erie, NY
Providence, RI
Beaver
Berks
Blair
Carbon
Juniata
Lackawanna
Lancaster
Luzerne
Schuykill
Westmoreland
York
PA
Page
Scott
VA
Alexander
Caldwell
Rockingham
Rutherford
Stokes
Surry
Wilkes
NC

Fig. IV-1. U.S. counties with total population greater than 10,000 in which white male in-migration and out-migration for 1965-1970 are both less than 12 percent of the 1970 county population total. (31 counties; none west of those shown here)

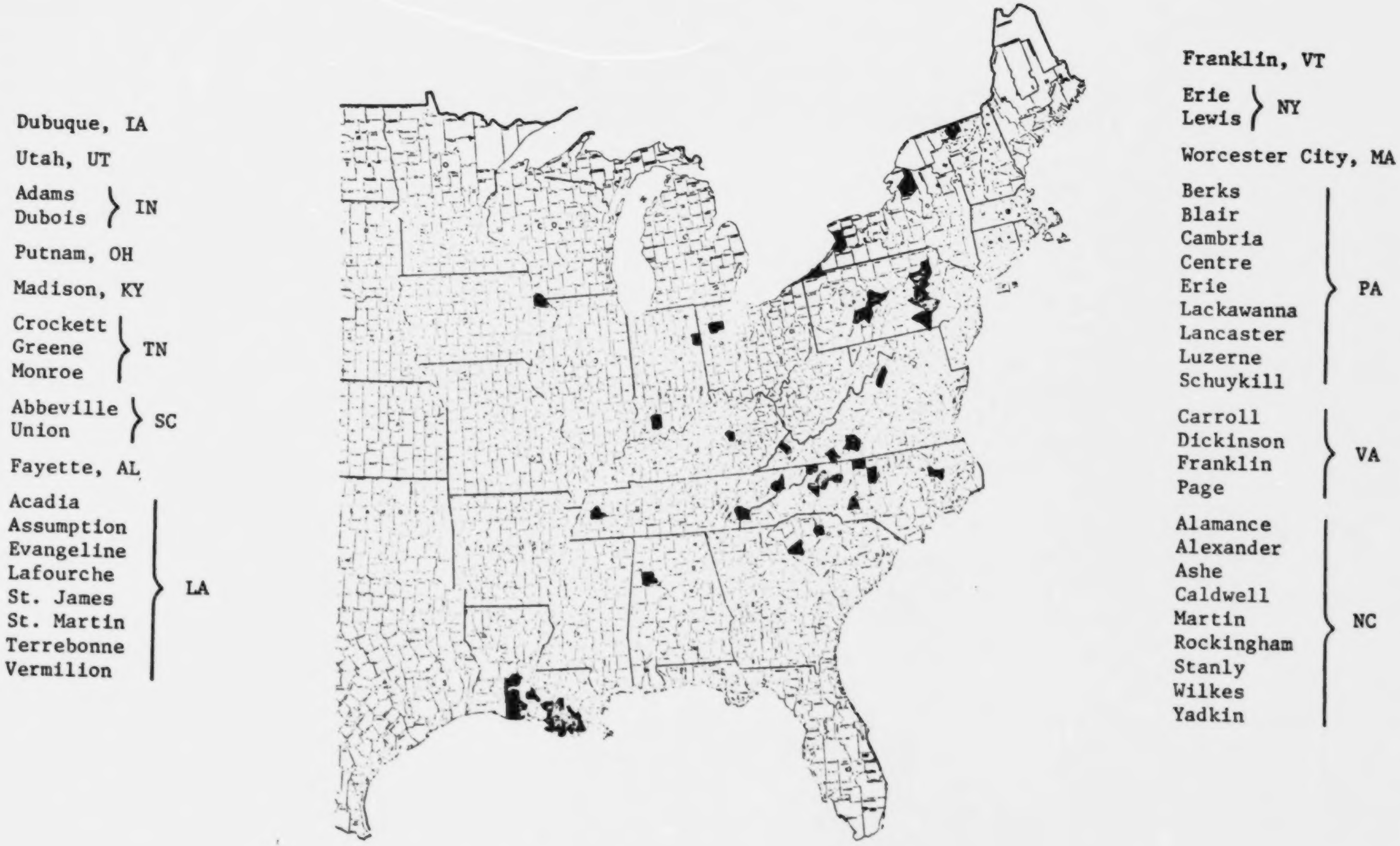

Dubuque, IA
Utah, UT
Adams
Dubois
IN
Putnam, OH
Madison, KY
Crockett
Greene
Monroe
TN
Abbeville
Union
SC
Fayette, AL
Acadia
Assumption
Evangeline
Lafourche
St. James
St. Martin
Terrebonne
Vermilion
LA
Franklin, VT
Erie
Lewis
NY
Worcester City, MA
Berks
Blair
Cambria
Centre
Erie
Lackawanna
Lancaster
Luzerne
Schuylkill
PA
Carroll
Dickinson
Franklin
Page
VA
Alamance
Alexander
Ashe
Caldwell
Martin
Rockingham
Stanly
Wilkes
Yadkin
NC

Fig. IV-2. U.S. counties (45 counties and Worcester city, MA) with total population greater than 10,000 in which white-male-over-the-age-of-44 in-migration and out-migration for 1965-1970 are both less than 1.75 percent of the 1970 county population total. (Utah County, UT, not shown)

TABLE IV-11

U.S. COUNTIES WITH TOTAL POPULATION GREATER THAN 10,000 IN WHICH THE RATIO OF IN-MIGRATION TO OUT-MIGRATION FOR WHITE MALES WHO WERE 45 YEARS OF AGE OR OLDER IN 1970 IS GREATER THAN 3.00

County	Total White Male In-Migration *	Total White Male Out-Migration *	WM 45+ In-Migration *	WM 45+ Out-Migration *	Ratio In/Out
ARIZONA					
Mohave	54.8	15.9	21.69	3.98	5.45
ARKANSAS					
Baxter	28.3	13.6	16.28	3.29	4.95
Cleburne	22.2	14.1	7.99	2.18	3.67
FLORIDA					
Broward	39.9	12.7	17.61	3.67	4.80
Charlotte	46.2	13.7	29.36	6.41	4.58
Citrus	47.7	11.5	27.54	4.33	6.36
Collier	46.5	14.2	21.07	3.02	6.98
Hernando	37.7	18.7	17.85	5.63	3.17
Highlands	36.3	25.9	17.42	5.73	3.04
Indian River	32.5	16.0	12.37	3.32	3.73
Lee	40.8	13.3	19.76	3.49	5.66
Levy	32.6	17.7	11.09	3.61	3.07
Manatee	32.8	14.3	17.84	3.86	4.62
Martin	41.1	14.9	17.91	4.72	3.79
Osceola	36.6	17.4	13.71	4.12	4.54
Palm Beach	34.3	14.3	13.79	3.04	4.54
Pasco	50.0	13.6	31.51	4.02	7.84
Pinellas	34.0	13.1	18.03	3.43	5.26
Sarasota	36.7	16.5	20.26	4.21	4.81
GEORGIA					
Jones	29.0	17.6	4.69	1.11	4.23
MICHIGAN					
Ogemaw	27.7	13.3	9.98	2.44	4.09

* % of County Census Population

TABLE IV-11 (continued)

U.S. COUNTIES WITH TOTAL POPULATION GREATER THAN 10,000 IN WHICH THE RATIO OF IN-MIGRATION TO OUT-MIGRATION FOR WHITE MALES WHO WERE 45 YEARS OF AGE OR OLDER IN 1970 IS GREATER THAN 3.00

County	Total White Male In-Migration *	Total White Male Out-Migration *	WM 45+ In-Migration *	WM 45+ Out-Migration *	Ratio In/Out
MISSISSIPPI					
De Soto	37.4	15.6	6.17	1.95	3.16
Jasper	17.7	13.7	3.51	0.90	3.90
NEW JERSEY					
Cape May	26.5	16.6	9.54	3.15	3.03
Ocean	32.6	14.7	12.56	2.70	4.65
TENNESSEE					
Rhea	13.9	12.6	2.65	0.76	3.49
Williamson	29.9	11.8	6.24	2.01	3.10
TEXAS					
Henderson	27.2	19.0	9.02	2.58	3.50
Mongtomery	42.5	13.9	9.08	2.19	4.15
VIRGINIA					
Westmoreland	20.6	12.9	8.32	2.33	3.57

* % of County Census Population

TABLE IV-12

U.S. COUNTIES WITH TOTAL POPULATION GREATER THAN 10,000 IN WHICH THE RATIO OF OUT-MIGRATION TO IN-MIGRATION FOR WHITE MALES WHO WERE 45 YEARS OF AGE OR OLDER IN 1970 IS GREATER THAN 2.00

County	Total White Male In-Migration*	Total White Male Out-Migration*	WM 45+ In-Migration*	WM 45+ Out-Migration*	Ratio Out/In
GEORGIA					
Cook	14.9	23.3	2.37	5.50	2.32
Grady	10.0	27.8	1.37	4.19	3.06
ILLINOIS					
Alexander	12.7	30.9	1.89	8.00	4.23
KENTUCKY					
Perry	8.1	24.4	1.58	3.91	2.47
MICHIGAN					
Wayne	10.7	21.4	1.60	5.05	3.16
MONTANA					
Big Horn	22.9	41.0	3.61	10.94	3.03
Dawson	18.2	36.2	1.87	5.37	2.87
NEBRASKA					
Cheyenne	18.3	38.1	2.36	8.10	3.43
NORTH CAROLINA					
Greene	11.2	22.3	1.00	4.22	4.22
Mitchell	7.5	17.7	0.83	3.09	5.19
VIRGINIA					
Buchanan	8.4	18.3	0.69	2.66	3.86

* % of County Census Population

Noting that 30 of the 75 counties listed had a white-male-over-the-age-of-64 population greater than 2,000 and that only in Florida, New Jersey and Pennsylvania did a majority of the listed counties satisfy this older population criterion, it was decided, in consultation with the EPA Project Director, that closer examination of Medicare statistics in ten Pennsylvania counties, 14 Florida counties, and one New Jersey county would be appropriate for further narrowing the field of candidate counties for use in the morbidity-mortality analysis to follow. The maps of Figs. IV-3, IV-4, and IV-5 show the 25 counties which were considered.

Enumerations of patient characteristics in the selected counties were obtained from the Medicare admission discharge records. In addition, population and industrial data were compiled for each of these counties. Findings are given in Tables IV-13, IV-14, and IV-15.

At a meeting with the EPA Project Director, the merits of the county choice alternatives were discussed and selection of the study areas (two Florida counties and a region in eastern Pennsylvania comprised of five contiguous counties) was made. Broward and Pinellas counties in Florida have the largest populations among the "high migration" counties; they offer an excellent opportunity to examine the potential of the Medicare data in estimating migrants' contribution to disease prevalence and specifying other characteristics of this relatively mobile population.

Berks, Lackawanna, Lancaster, Luzerne, and Schuylkill counties in Pennsylvania were also selected for analysis. They are among the most "stable" (low migration) counties nationwide, and, by forming a contiguous region, offer an opportunity to contrast patterns and causes of hospital admissions among the Medicare-eligible sample population over a fairly constrained and demographically homogeneous area.

4. Industrial Development Patterns in the Selected Counties

Following designation of the study areas, documentation of industrial development in the seven counties (two in Florida, five in Pennsylvania) was

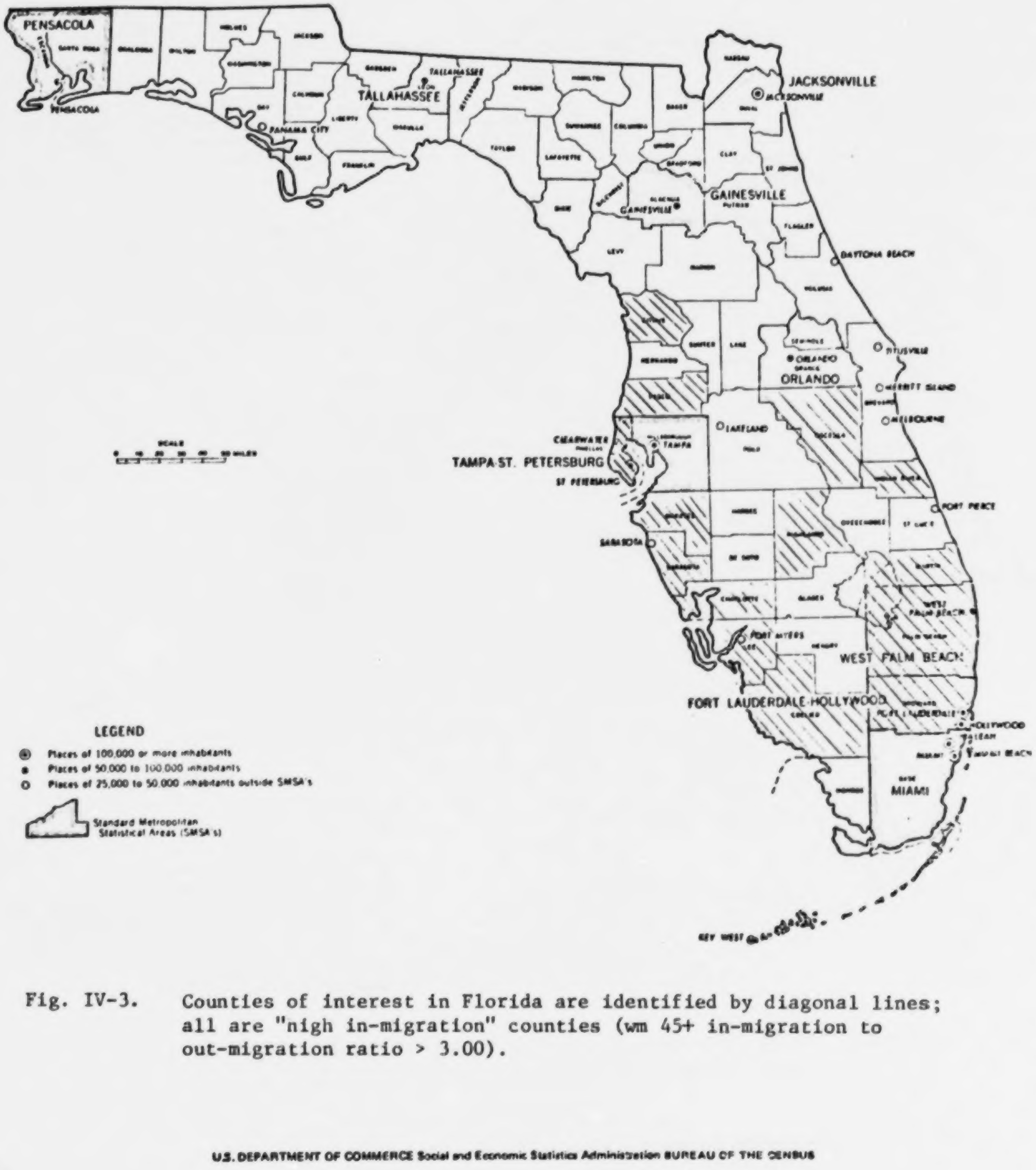

PENSACOLA
TALLAHASSEE
PANAMA CITY
JACKSONVILLE
GAINESVILLE
DAYTONA BEACH
TITUSVILLE
ORLANDO
MELBOURNE
LAKELAND
CLEARWATER
TAMPA
TAMPA-ST. PETERSBURG
ST. PETERSBURG
SARASOTA
FORT PIERCE
FORT MYERS
WEST PALM BEACH
FORT LAUDERDALE-HOLLYWOOD
HOLLYWOOD
MIAMI BEACH
MIAMI
KEY WEST
SCALE
LEGEND
Places of 100,000 or more inhabitants
Places of 50,000 to 100,000 inhabitants
Places of 25,000 to 50,000 inhabitants outside SMSA's
Standard Metropolitan Statistical Areas (SMSA's)

Fig. IV-3. Counties of interest in Florida are identified by diagonal lines; all are "nigh in-migration" counties (wm 45+ in-migration to out-migration ratio > 3.00).

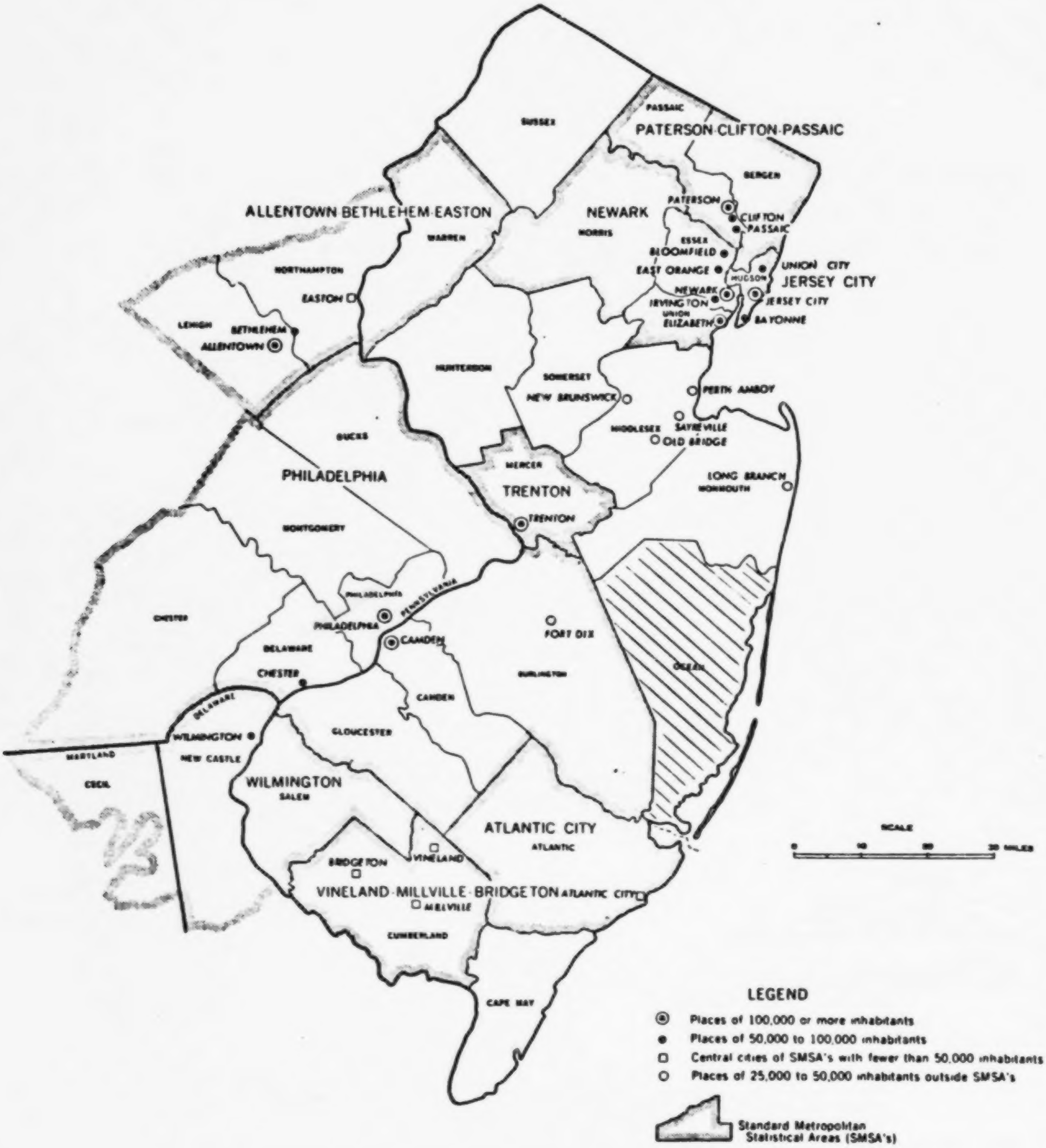

SUSSEX
PASSAIC
PATERSON-CLIFTON-PASSAIC
BERGEN
ALLENTOWN-BETHLEHEM-EASTON
NEWARK
WARREN
MORRIS
PATERSON
CLIFTON
PASSAIC
ESSEX
BLOOMFIELD
EAST ORANGE
UNION CITY
HUDSON
JERSEY CITY
NORTHAMPTON
EASTON
NEWARK
IRVINGTON
JERSEY CITY
UNION
ELIZABETH
BAYONNE
LEHIGH
BETHLEHEM
ALLENTOWN
HUNTERDON
SOMERSET
PERTH AMBOY
NEW BRUNSWICK
MIDDLESEX
SAYREVILLE
OLD BRIDGE
BUCKS
PHILADELPHIA
MERCER
TRENTON
LONG BRANCH
MONMOUTH
MONTGOMERY
TRENTON
PHILADELPHIA
PENNSYLVANIA
CHESTER
PHILADELPHIA
FORT DIX
CAMDEN
DELAWARE
OCEAN
CHESTER
BURLINGTON
CAMDEN
DELAWARE
WILMINGTON
GLOUCESTER
MARYLAND
NEW CASTLE
CECIL
WILMINGTON
SALEM
ATLANTIC CITY
ATLANTIC
SCALE
0
10
20
30 MILES
VINELAND
BRIDGETON
VINELAND-MILLVILLE-BRIDGETON
ATLANTIC CITY
MILLVILLE
CUMBERLAND
CAPE MAY
LEGEND
◉ Places of 100,000 or more inhabitants
● Places of 50,000 to 100,000 inhabitants
□ Central cities of SMSA's with fewer than 50,000 inhabitants
○ Places of 25,000 to 50,000 inhabitants outside SMSA's
Standard Metropolitan Statistical Areas (SMSA's)

Fig. IV-4. County of interest in New Jersey is identified by diagonal lines; Ocean County is a "high in-migration" county.

Reproduced from
best available copy.

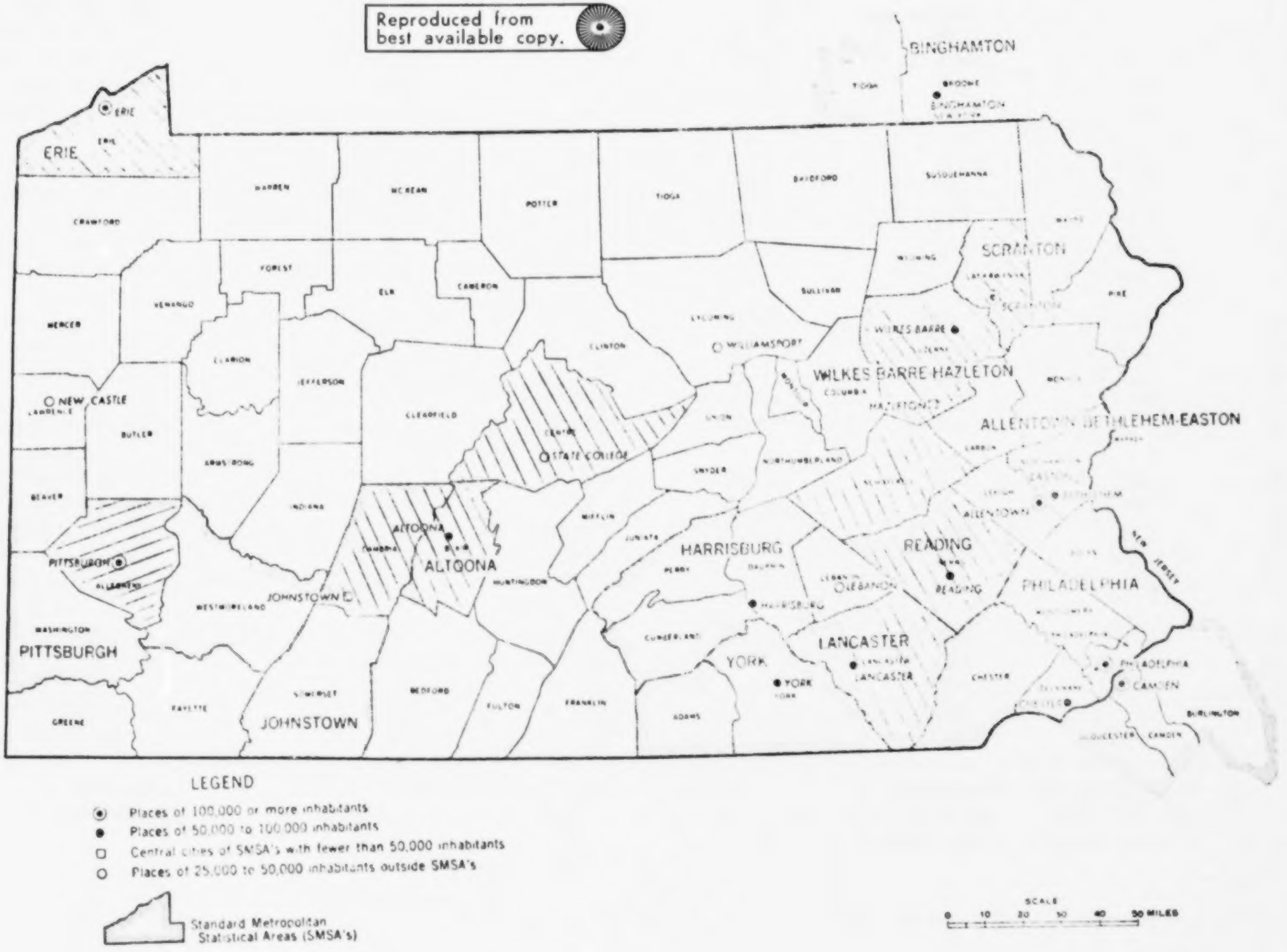

BINGHAMTON
BROOME
BINGHAMTON
TIOGA
ERIE
ERIE
CRAWFORD
WARREN
MC KEAN
POTTER
TIOGA
BRADFORD
SUSQUEHANNA
WAYNE
SCRANTON
LACKAWANNA
SCRANTON
PIKE
WYOMING
SULLIVAN
LYCOMING
WILLIAMSPORT
WILKES BARRE
LUZERNE
WILKES BARRE HAZLETON
COLUMBIA
HAZLETON
MONROE
ALLENTOWN BETHLEHEM EASTON
CARBON
EASTON
BETHLEHEM
ALLENTOWN
LEHIGH
SCHUYLKILL
NORTHUMBERLAND
UNION
SNYDER
MIFFLIN
JUNIATA
CLINTON
CENTRE
STATE COLLEGE
FOREST
ELK
CAMERON
MERCER
VENANGO
CLARION
JEFFERSON
CLEARFIELD
NEW CASTLE
LAWRENCE
BUTLER
ARMSTRONG
INDIANA
BEAVER
PITTSBURGH
ALLEGHENY
WESTMORELAND
JOHNSTOWN
CAMBRIA
ALTOONA
BLAIR
ALTOONA
HUNTINGDON
WASHINGTON
PITTSBURGH
GREENE
FAYETTE
SOMERSET
JOHNSTOWN
BEDFORD
FULTON
FRANKLIN
HARRISBURG
PERRY
DAUPHIN
HARRISBURG
CUMBERLAND
ADAMS
YORK
YORK
YORK
LEBANON
LEBANON
LANCASTER
LANCASTER
LANCASTER
READING
BERKS
READING
PHILADELPHIA
BUCKS
MONTGOMERY
PHILADELPHIA
CHESTER
DELAWARE
CHESTER
PHILADELPHIA
CAMDEN
BURLINGTON
GLOUCESTER
CAMDEN
NEW JERSEY
LEGEND
⊙ Places of 100,000 or more inhabitants
● Places of 50,000 to 100,000 inhabitants
□ Central cities of SMSA's with fewer than 50,000 inhabitants
○ Places of 25,000 to 50,000 inhabitants outside SMSA's
Standard Metropolitan Statistical Areas (SMSA's)
SCALE
0 10 20 30 40 50 MILES

Fig. IV-5. Counties of interest in Pennsylvania are diagonal lined; all but Allegheny County are "stable" (wm 45+ in- and out-migration < 1.75%).

TABLE IV-13
COUNTY POPULATION CHARACTERISTICS

Selected Counties	Total Population	White Males Over Age 64	% Of White Male Population	White Male Over Age 64 In Migration	White Male Over Age 64 Out Migration	% Of White Male Population Over Age 64 In/Out
Broward FL	620,100	51,471	20.0	20,950	3,804	40.7 / 7.4
Charlotte FL	27,559	4,999	39.1	2,150	468	43.0 / 9.4
Citrus FL	19,196	2,537	30.2	1,169	112	46.1 / 4.4
Collier FL	38,040	2,740	15.9	1,289	140	47.0 / 5.1
Highlands FL	29,507	2,726	24.3	1,051	202	38.6 / 7.4
Indian River FL	35,992	2,739	19.5	645	170	23.6 / 6.2
Lee FL	105,216	9,643	21.5	3,954	584	41.0 / 6.1
Manatee FL	97,115	13,289	33.4	4,147	772	31.2 / 5.8
Martin FL	28,035	2,857	24.6	957	244	33.5 / 8.5
Osceola FL	25,267	2,444	22.2	673	192	27.5 / 7.9
Palm Beach FL	348,753	26,121	19.2	9,144	1,497	35.0 / 5.7
Pasco FL	75,955	12,319	35.4	6,622	707	53.8 / 5.7
Pinellas FL	522,329	67,212	30.6	21,991	3,537	32.7 / 5.3
Sarasota FL	120,413	16,114	30.8	5,744	1,050	35.7 / 6.5
Ocean NJ	208,470	15,219	15.6	6,186	995	40.7 / 6.5
Allegheny PA	1,605,016	65,776	9.5	2,146	4,777	3.3 / 7.3
Berks PA	296,382	14,663	10.6	721	680	4.9 / 4.6
Blair PA	135,356	7,023	11.1	196	298	2.8 / 4.2
Cambria PA	186,785	8,676	9.8	361	414	4.2 / 4.8
Centre PA	99,267	2,674	5.3	198	178	7.4 / 6.7
Erie PA	263,654	10,086	8.2	527	686	5.2 / 6.8
Lackawanna PA	234,107	12,088	11.1	578	469	4.8 / 3.9
Lancaster PA	319,693	13,306	8.8	699	471	5.3 / 3.5
Luzerne PA	342,301	17,425	10.9	656	592	3.8 / 3.4
Schuylkill PA	160,089	8,731	11.5	228	342	2.6 / 3.9

TABLE IV-14

MEDICARE DATA CHARACTERISTICS

Selected Counties	White Male Medicare Claimants In 1971-1973 Sample	White Male Cancer Patients In Sample	% Of White Male Claimants	White Male Respiratory Cancer Patients In Sample	% Of White Male Cancer Patients
Broward FL	2,418	931	38.5	174	18.7
Charlotte FL	219	57	26.0	11	19.3
Citrus FL	129	31	24.0	5	16.1
Collier FL	145	55	37.9	16	29.1
Highlands FL	142	46	32.4	11	23.9
Indian River FL	159	61	38.4	8	13.1
Lee FL	453	147	32.5	22	15.0
Manatee FL	486	164	33.7	27	16.5
Martin FL	127	53	41.7	9	17.0
Osceola FL	138	42	30.4	8	19.0
Palm Beach FL	1,229	438	35.6	86	19.6
Pasco FL	409	150	36.7	24	16.0
Pinellas FL	2,660	935	35.2	167	17.9
Sarasota FL	690	259	37.5	52	20.1
Ocean NJ	719	267	37.1	61	22.8
Allegheny PA	2,872	1,008	35.1	176	17.5
Berks PA	566	212	37.5	25	11.8
Blair PA	264	59	22.3	7	11.9
Cambria PA	412	111	26.9	18	16.2
Centre PA	123	40	32.5	4	10.0
Erie PA	432	130	30.1	7	5.4
Lackawanna PA	517	165	31.9	19	11.5
Lancaster PA	534	209	39.1	23	11.0
Luzerne PA	825	220	26.7	38	17.3
Schuylkill PA	463	120	25.9	22	18.3

TABLE IV-15

COUNTY URBAN/INDUSTRIAL CHARACTERISTICS

Selected Counties	% Of Population Urban	Number of Manufacturing Establishments	% With 20-99 / 100+ Employees/Employees	Annual Average Number Of Employees In Manufacture	% Of Total Population
Broward FL	99.0	680	19.1/ 4.6	15,800	2.5
Charlotte FL	59.1	19	15.8/ -	200	0.7
Citrus FL	0.0	16	6.3/ -	100	0.5
Collier FL	66.1	25	20.0/ -	300	0.8
Highlands FL	47.2	23	17.4/ -	300	1.0
Indian River	69.6	33	21.2/ 6.1	2,500	6.9
Lee FL	70.3	88	23.9/ 1.1	1,400	1.3
Manatee FL	71.4	79	19.0/10.1	3,300	3.4
Martin FL	16.8	28	14.3/ -	400	1.4
Osceola FL	47.6	35	25.7/ -	700	2.8
Palm Beach FL	91.1	324	18.8/ 5.6	15,400	4.4
Pasco FL	33.8	44	9.1/ 4.5	1,400	1.8
Pinellas FL	96.1	511	17.4/ 6.1	18,400	3.5
Sarasota FL	75.0	161	14.3/ 1.9	2,700	2.2
Ocean NJ	44.3	184	18.5/ 2.7	3,800	1.8
Allegheny PA	94.8	1,663	26.3/14.3	195,400	12.2
Berks PA	63.5	666	31.7/18.8	56,400	19.0
Blair PA	67.3	158	21.5/21.5	14,600	10.8
Cambria PA	59.0	182	27.5/11.0	21,500	11.5
Centre PA	47.2	101	22.8/15.8	7,400	7.5
Erie PA	74.8	461	25.6/18.9	42,300	16.0
Lackawanna PA	87.2	482	38.8/18.5	33,700	14.4
Lancaster PA	54.2	660	29.2/16.5	54,100	16.9
Luzerne PA	78.2	778	40.7/15.6	49,700	14.5
Schuylkill PA	51.9	313	43.8/19.2	21,700	13.6

undertaken. The number of employees in each of several distinct industrial activities was used as the measure of industrialization. Employment data from County Business Patterns for the years 1953, 1959, 1965, 1970, and 1974 were obtained for each of the seven counties at the two-digit level of Standard Industrial Classification (SIC) code, except for industry codes identified by EPA's Implementation Planning Program as relating to significant sources of TSP or SO_2. For the latter industry groups, data were obtained for the most detailed code available.

A perspective of chronological changes in total employment and the main types of employment in the selected counties is presented in Fig. IV-6A, for the Florida counties, and Fig. IV-6B, for the counties in Pennsylvania. The categories of primary metal industries, and chemical, petroleum, and rubber displayed in the bar charts of Fig. IV-6 were chosen to facilitate comparison with the example study of steel-industry counties, to be described later.

The directors of the statistical divisions of the Pennsylvania and Florida Departments of Commerce also supplied data maintained by their respective offices concerning the chronological development of industrial employment within the study counties. The Bureau of Statistics, Research, and Planning of the Pennsylvania Department of Commerce forwarded Industrial Census Reports for the five counties of interest for each of the years 1970-1976. In addition to historical trends of employment, production, and wages, there is a complete four-digit SIC categorization of establishments, with a distinction between "production (and related) workers" and "all others." These reports were first printed in this format in 1961. Prior year data (back to 1950) were included in Productive Industries Reports. The Division of Employment Security of the Florida Department of Commerce sent information on historical employment data (including sample reports and descriptions of data limitation) for Broward and Pinellas counties. Most data were provided in a two- or three-digit SIC code. However, the data from County Business Patterns, Bureau of Census, were more generally useful for consistency, uniformity, and the demonstration purposes of this project.

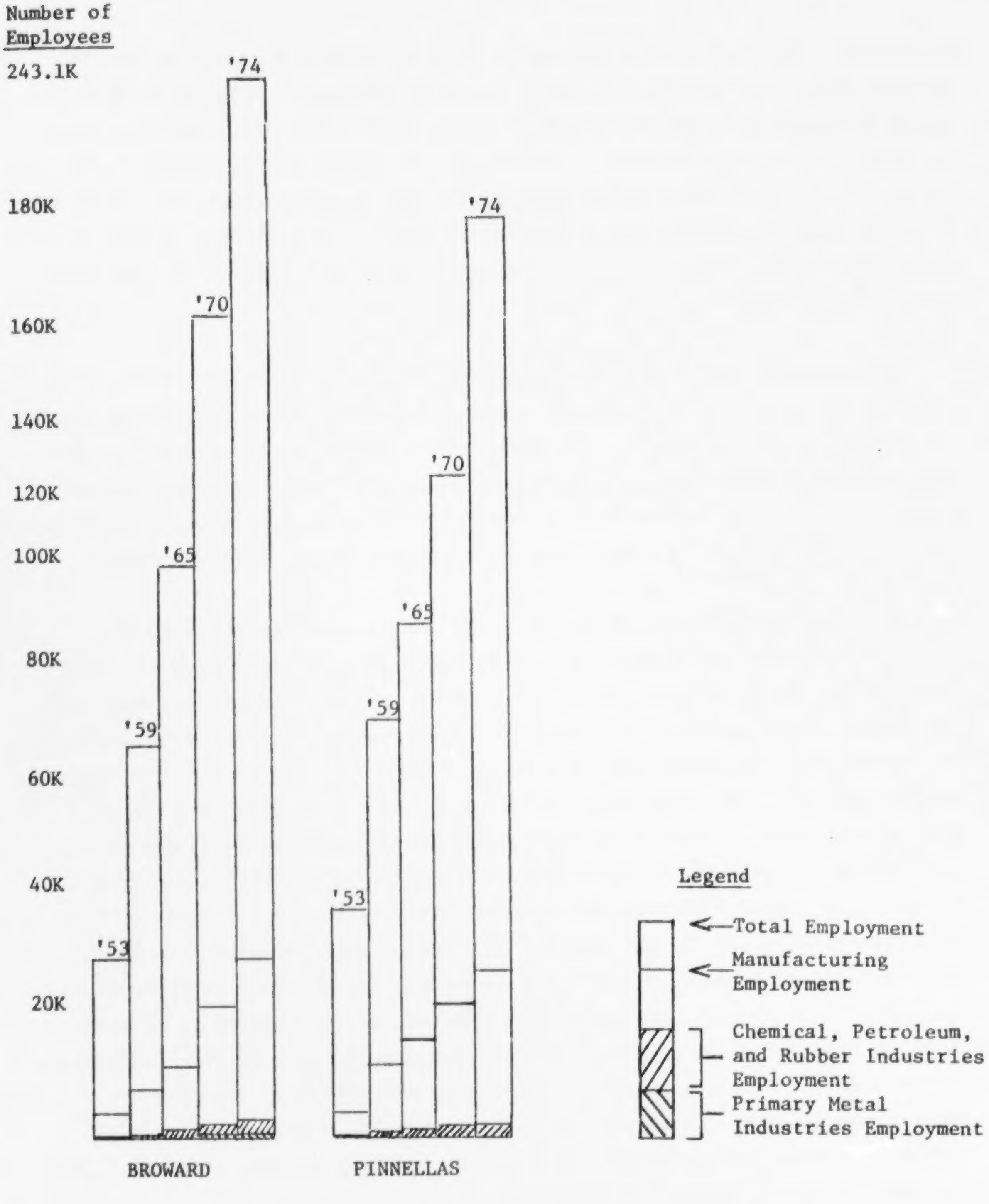


Fig. IV-6A. Levels of Employment in Two Florida Counties for the Years 1953, 1959, 1965, 1970 and 1974

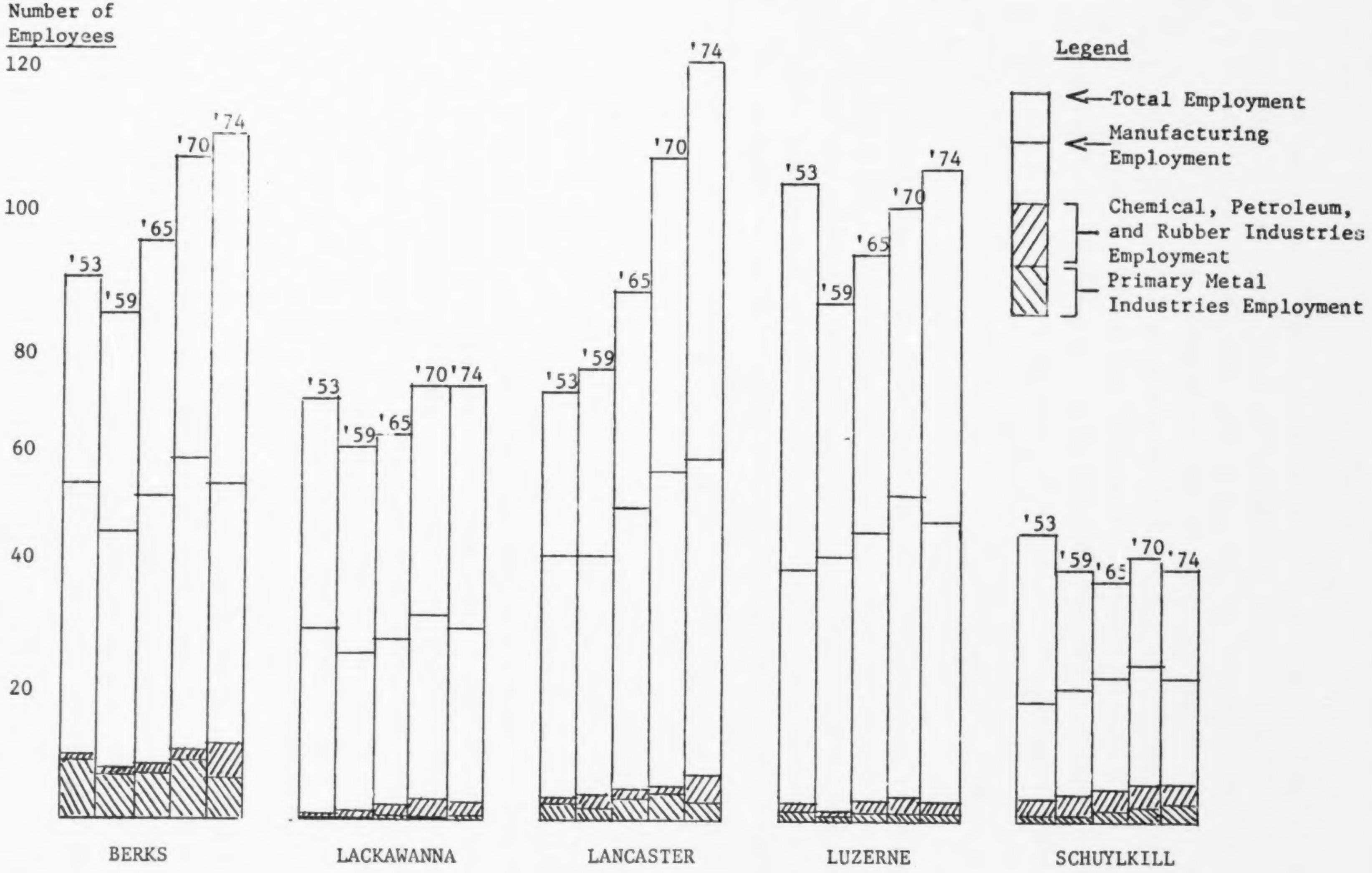


Fig. IV-6B. Levels of Employment in Five Pennsylvania Counties for the Years 1953, 1959, 1965, 1970 and 1974.

A particularly interesting publication received was "Occupational Employment in Manufacturing Industries in Florida," in which estimates of total and proportional numbers of employees within occupational subdivisions of the three-digit SIC's 201-399 were derived from a 1974 survey of 5,550 firms in the state. This information was obtained for the first time in 1974. The survey will be repeated at three-year intervals, and holds considerable promise as an analytical tool for epidemiological studies of Health Effects Research Laboratory, EPA, in linking industries with occupations.

5. Morbidity/Mortality Comparisons

A fundamental question being addressed by this study is whether the morbidity data collected as part of the Medicare program adequately reflects--or, indeed, extends--knowledge of disease incidence and prevalence that can otherwise be obtained from mortality statistics. This section presents three illustrations of the manner in which the Medicare records can be retrieved for purposes of health effects investigations; each example focuses on a different geographical area and a different time interval.

A brief review of mortality information is helpful for an understanding of the complementary role of morbidity information. The National Center for Health Statistics receives copies of death certificates from vital statistics offices in each of the 50 states and U.S. territories, then codes much of the non-confidential information for computer storage, research purposes, and the preparation of summary statistics. Mortality rates computed from these data have, to this time, been based almost exclusively on the single code for the "underlying cause of death." Items on the death certificate form pertaining to other causal conditions manifest at the time of death are not coded for computer-assisted analyses.

The significance of use of the "underlying cause" alone can be appreciated by reference to a recent study of the 1969 death certifications for the United States. In that year, malignant neoplasms of some kind were "mentioned" over 565,000 times as contributory, associated and/or the underlying cause of death on the certificates of some 368,000 individuals. Of these, 323,000 died with some malignant neoplasm designated as the underlying cause; customary epidemiological analyses would be limited to these 323,000 cases. Within the neoplasm total were 15,989 "mentions" of secondary malignant neoplasms of the lung, with 766 of them designated as the underlying cause of death. Use of the underlying cause alone would ignore over 15,000 cases, and significantly understates the prevalence of secondary malignant neoplasms of the lung. Lung diseases are particularly important to EPA and its concern with human health effects.

Two ways of obtaining more complete information concerning the numbers of people affected by certain diseases are (1) use of all medical descriptors entered on the death certification form and (2) use of medical histories and/or hospitalization records. The Medicare file is an example of this second alternative. These two ways are not mutually exclusive. They are complementary. With complex interactions of the real world, neither the mortality nor the morbidity records can be expected to meet fully the needs of the environmental health analyst.

There are advantages and limitations to each of the two types of records. The death certificate, especially when augmented by an autopsy, is generally complete and accurate; but is restricted to observations concerning the death. Medical histories, as represented by the Medicare hospitalization sample, often provide medical information well before the time of death, including possible chronological development of disease conditions. Medical histories also give some indication of geographical location which may be related to pollution levels, or time-specific events detrimental to health. Unfortunately, there is no source of medical history data that meets all environmental health analysis needs. The Medicare files utilized are more accessible and extensive than other medical-history sources, but nevertheless the files pertain only to the older members of the population and are a 20 percent sample. At present, the Medicare data has only one disease code with each hospital admission/discharge--a limitation identical with mortality data.

Nonetheless, the Medicare data obtained from the retrievals described below do give an important perspective on the problem of identifying the number of individuals within each county recognized each year as having certain diseases.

Morbidity trends are obviously more timely than mortality trends, and offer the singular advantage of possible interventions before death results. As a further practical consideration, there are about 16 times more hospital admissions than deaths.

It is noted that the POPATRISK data base was not used as the source of mortality data for mortality/morbidity comparisons because POPATRISK contains age-adjusted, rather than age-specific, rates and would not be comparable to the simple counts obtained from the Medicare record files. For convenience, the mortality data used were provided directly by the State of Pennsylvania.

a. Berks County--By Year--

The first morbidity/mortality comparison example is for one county, Berks, in Pennsylvania. A single county was chosen because of the wealth of demographic, environmental pollution, economic and general statistical information that is collected and coded for computer-assisted analyses by standard county identifiers. With these identifiers and codes, widely-used computer procedures can be directed to inter-county comparisons, as well as assembling the data for comparisons by metropolitan areas and states.

The Pennsylvania Department of Health compiled mortality data for Berks, Lackawanna, Lancaster, Luzerne, and Schuylkill counties according to disease categories specified by System Sciences, Inc., identical to the 19 cancer diagnosis subdivisions presented in Table IV-9. There are five-year age groups through age 74, also 75-84 and 85+. Statistics in all seven counties cover the individual years 1971-1975, as requested. For pilot analysis purposes, data were aggregated, without differentiating by sex or race of the deceased. The mortality statistics thus obtained provide an aggregate basis for comparison with the Medicare-derived statistics.

The retrievals from the 1971-1973 Medicare files counted hospitalizations for residents of the seven selected study counties, including Berks. Counts by county, year, disease category and age were taken of the following: the number of hospitalizations in which the claimant was discharged dead and the number of new cases of cancer each year partitioned by the year of death of the claimant or the fact that the death of the claimant is not indicated. Morbidity and mortality data for Berks County only are presented and discussed at this time.

The cancer mortality statistics provided by the Pennsylvania Department of Health for Berks County for the years 1971, 1972, and 1973 are shown in Table IV-16 under the four age groups 65-69, 70-74, 75-84, and over 84 years with 19 subdivisions of the neoplasm disease classification. The 92 deaths over the three-year period (30, 24, and 38 deaths in the respective years) give a small number of cases in each of the age-year-disease categories; little can be determined from these data with regard to mortality trends.

Table IV-17 consolidates the age categories for the mortality statistics from both the Medicare file and the Pennsylvania vital statistics office; the counts are by year and cancer type for Berks County residents over the age of 64. While there is an apparent upward trend from 1971 to 1973 in the number of hospital deaths reported for cancer of the trachea/bronchus/lung (ICDA 162) and the miscellaneous classification (190, 192-199), there is undoubtedly a high degree of variability in the sample data, and there is not a comparable indication of trend from the statistics of the Pennsylvania Department of Health. It can be seen in Table IV-17 that the deaths reported by Medicare for Berks County are not a uniform proportion of the number of deaths as reported by the Commonwealth of Pennsylvania.

The changes in the number of cases in the specified disease categories during each year are presented in Table IV-18. A table for each of the years 1971, 1972, 1973 shows the net change in cases for each disease category; that is, the number of new cases first appearing in the years, minus the cases removed by death during the same year.

Table IV-18 also shows the net change over the entire three-year period covered by the consolidated Medicare files available for this project, and

TABLE IV-16

DEATH COUNTS BY NEOPLASM TYPE AND AGE GROUP FOR BERKS COUNTY, PA AND YEARS 1971, 1972, AND 1973

Neoplasms	Ages 65-69			Ages 70-74			Ages 75-84			Ages 85+		
	1971	1972	1973	1971	1972	1973	1971	1972	1973	1971	1972	1973
140-145, buccal cavity	0	0	0	0	0	0	[illegible]	0	0	0	0	0
146-149, pharynx	0	0	0	0	0	0	0	0	0	0	0	0
151, stomach	0	0	0	0	1	1	0	0	0	0	0	0
153, large intestine (exc. rectum)	1	0	0	2	1	1	2	0	1	0	0	1
154, rectum/rectosigmoid junction	1	0	0	0	0	0	0	0	0	1	0	1
157, pancreas	0	0	0	0	0	0	0	0	0	0	0	0
150, 152, 155, 156, 158, 159, other digestive	1	0	0	1	0	2	0	0	1	0	0	0
162, trachea/bronchus/lung	2	1	3	0	1	2	0	0	3	0	1	0
160, 161, 163, other respiratory	0	0	0	0	0	0	0	0	1	0	0	0
170-174, bone/connective tissue/ skin/breast	0	0	1	1	1	2	0	0	1	0	0	0
185, prostate	0	0	1	0	0	1	2	2	0	0	0	1
188, bladder	0	0	0	1	0	0	1	0	0	2	1	0
180-184, 186, 187, 189, other genitourinary	2	0	0	0	0	1	0	0	0	0	1	0
191, brain	0	0	0	0	0	0	0	0	0	0	0	0
190, 192-199, other and unspecified	2	2	2	1	4	3	0	1	1	1	0	3
204-207, leukemia	1	0	0	0	2	0	1	0	1	0	0	0
200-203, 208, 209, other lymphatic and hematopoietic tissue	1	1	1	0	3	0	1	1	2	0	0	0
210-228, benign	0	0	0	0	0	0	0	0	0	0	0	0
230-239, unspecified nature	1	0	0	0	0	0	0	0	0	0	0	0
Total all Neoplasms	12	4	8	6	13	13	8	4	11	4	3	6

Source: Pennsylvania Department of Health.

TABLE IV-17

DEATH COUNTS BY NEOPLASM TYPE FOR PERSONS AGED 65 OR OVER IN BERKS COUNTY, PA

ICDA Code	Number Discharged Dead with Indicated Neoplasm Type as Discharge Diagnosis (obtained from Medicare Hospitalization Sample)			Number with Indicated Neoplasm Type Reported as Cause of Death to Pennsylvania Department of Health		
	1971	1972	1973	1971	1972	1973
Neoplasms						
140-145, buccal cavity	1	0	0	2	2	2
146-149, pharynx	0	0	0	1	2	0
151, stomach	0	1	1	22	17	13
153, large intestine (exc. rectum)	5	1	3	47	38	52
154, rectum/rectosigmoid junction	2	0	1	20	15	10
157, pancreas	0	0	0	18	24	13
150, 152, 155, 156, 158, 159, other digestive	2	0	3	22	10	17
162, trachea/bronchus/lung	2	3	8	50	59	54
160, 161, 163, other respiratory	0	0	1	6	2	7
170-174, bone/connective tissue/ skin/breast	1	1	4	29	31	28
185, prostate	2	2	3	29	26	33
188, bladder	4	1	0	15	13	16
180-184, 186, 187, 189, other genitourinary	2	1	1	29	25	22
191, brain	0	0	0	9	1	2
190, 192-199, other and unspecified	4	7	9	30	47	42
204-207, leukemia	2	2	1	11	13	13
200-203, 208, 209, other lymphatic and hematopoietic tissue	2	5	3	14	27	15
210-228, benign	0	0	0	0	1	0
230-239, unspecified nature	1	0	0	3	6	4
Total all Neoplasms	30	24	38	357	362	343

TABLE IV-18

USE OF MEDICARE DATA TO ESTIMATE CANCER PREVALENCE

(Berks County 1971, 1972, 1973)

Hospital Diagnosis ICD Type/Group	Net Annual increase in Surviving Cases*			Cumulative Prevalence/ Survivors*
	1971	1972	1973	1973
140-145X, buccal cavity	3	0	0	3
146-149X, pharynx	1	2	0	3
151-151X, stomach	4	1	4	9
153-153X, large intestine (exc. rectum)	13	19	16	48
154-154X, rectum/rectosigmoid junction	5	5	9	19
157-157X, pancreas	2	2	0	4
150, 152, 155, 156, 158, 159, other digestive	1	1	2	4
162-162X, trachea/bronchus/lung	6	6	5	17
160, 161, 163X, other respiratory	6	1	0	7
170-174X, bone/connective tissue/skin/breast	20	18	16	54
185-185X, prostate	16	8	8	32
188-188X, bladder	9	8	7	24
180-184. 186, 187, 189X, other genitourinary	7	5	9	21
191-191X, brain	0	0	1	1
190, 192-199X, other and unspecified	7	10	11	28
204-207X, leukemia	0	3	2	5
200-203, 208, 209, other lymphatic and hematopoietic tissue	5	2	1	8
210-228X, benign	21	13	15	49
230-239X, unspecified nature	16	13	5	34

---

*Excludes those known to have died, each year, regardless of cause of death. For example, a 1971 lung neoplasm diagnosis (ICD 162), who dies in 1971, 1972, or 1973 of heart disease or any other cause is not classed as a survivor in the year in which death occurs.

represents all disease cases for Berks County Medicare claimants who were alive through 1973. Table IV-18 demonstrates the potential use of Medicare data for estimating the increasing and/or decreasing population burden of diseases of special interest to EPA on a county-by-county basis. The annual data illustrate the availability of equally important trend data. Trends, together with increasing/decreasing population with certain diagnoses, in and out of hospitals, present a capability believed to have important applications in studies of environmentally-related morbidity. The morbidity data resource may also prove invaluable as an indicator of the health improvements resulting from pollution abatement and control requirements of EPA. To have sufficiently large numbers for statistical analyses, however, it may be necessary to aggregate age and/or disease categories.

b. Morbidity/Mortality During the 1973 Allegheny County Air Pollution Episode--

The relevance of the Medicare data with respect to air pollution episodes was demonstrated by an application centered on a specific episode--Allegheny County in 1973. This episode was selected because of previous day-by-day mortality studies by EPA, and the readily available mortality data. It will be recalled that the Medicare files show day of admission as well as disease cause. Retrieval and analytical steps can thus be keyed directly to the days during or after an episode and to any area of interest in the United States consisting of one or more counties.

The EPA study of health effects related to an air pollution episode in Allegheny County, Pennsylvania, from August 26 through September 7, 1973. Findings revealed that the total number of county-resident deaths (all causes) for each of the 13 days from August 26 through September 7 exceeded the corresponding 1971-1973 daily mortality averages.

Table IV-19 displays the daily mortality totals for Allegheny County from August 1 through September 22 for the years 1971, 1972, and 1973. Deaths with either respiratory or cardiovascular disease as the underlying cause are also presented in this same table. These latter data show a greater variability,

TABLE IV-19

ALLEGHENY COUNTY, PA. RESIDENTS WHO DIED IN THE COUNTY, AUGUST 1 THROUGH SEPTEMBER 22, 1971-1973, WITH SUBTOTALS FOR UNDERLYING DEATH CAUSE OF RESPIRATORY OR CARDIOVASCULAR DISEASES

DATE	NO. DEAD IN 1971			DATE	NO. DEAD IN 1972			DATE		NO. DEAD IN 1973		
	TOTAL	RESP	C.V.		TOTAL	RESP	C.V.			TOTAL	RESP	C.V.
710801	38	2	17	720801	47	2	19	730801		51	6	24
710802	42	1	27	720802	43	1	19	730802		37	0	22
710803	42	0	26	720803	37	2	17	730803		38	0	22
710804	34	1	21	720804	31	2	16	730804		36	1	20
710805	44	3	23	720805	35	1	17	730805		44	4	19
710806	36	1	16	720806	37	2	18	730806		44	1	24
710807	46	1	24	720807	44	0	27	730807		41	2	24
710808	55	1	25	720808	39	1	25	730808		45	1	26
710809	40	2	19	720809	38	1	22	730809		44	1	26
710810	38	1	16	720810	31	0	18	730810		36	0	16
710811	39	0	26	720811	33	2	14	730811		43	0	20
710812	35	3	16	720812	31	1	15	730812		33	2	15
710813	51	5	21	720813	42	2	23	730813		45	3	24
710814	52	1	26	720814	48	3	30	730814		25	3	12
710815	42	3	16	720815	30	3	16	730815		50	2	25
710816	36	2	25	720816	40	2	26	730816		42	0	23
710817	45	4	19	720817	35	1	19	730817		41	0	28
710818	42	1	25	720818	50	5	30	730818		31	1	14
710819	56	3	32	720819	38	3	21	730819		34	2	17
710820	39	0	19	720820	38	0	25	730820		38	0	25
710821	51	1	28	720821	52	1	25	730821		55	1	30
710822	32	5	17	720822	34	1	24	730822		47	2	25
710823	34	0	21	720823	40	2	22	730823		38	0	20
710824	44	1	19	720824	45	2	32	730824		40	1	24
710825	41	7	17	720825	64	1	34	730825		37	2	17
710826	36	1	20	720826	37	3	20	730826		52	2	29
710827	34	0	16	720827	37	5	17	730827	E	50	3	28
710828	32	6	14	720828	44	1	26	730828	P	59	6	20
710829	34	1	18	720829	39	2	17	730829	I	67	7	33
710830	35	0	25	720830	44	2	24	730830	S	51	2	26
710831	42	5	20	720831	41	1	25	730831	O	53	3	33
710901	35	2	18	720901	26	1	11	730901	D	58	5	32
710902	31	0	16	720902	44	5	27	730902	E	57	7	29
710903	32	1	20	720903	29	0	11	730903		44	3	23
710904	49	0	29	720904	20	0	11	730904		50	3	33
710905	56	2	38	720905	42	1	23	730905		56	6	31
710906	50	2	26	720906	39	2	20	730906		49	5	23
710907	41	3	23	720907	37	4	24	730907		46	6	19
710908	52	1	29	720908	28	1	11	730908		37	2	16
710909	42	1	22	720909	40	3	19	730909		38	0	22
710910	39	1	21	720910	33	1	17	730910		44	0	25
710911	44	2	22	720911	45	0	26	730911		43	0	25
710912	33	3	19	720912	37	1	20	730912		37	1	19
710913	35	1	19	720913	50	0	26	730913		34	2	19
710914	26	1	18	720914	34	1	17	730914		43	2	21
710915	37	3	19	720915	38	2	23	730915		49	0	27
710916	40	2	18	720916	44	4	24	730916		41	4	21
710917	52	1	30	720917	43	1	24	730917		26	1	16
710918	35	1	17	720918	50	3	33	730918		50	4	25
710919	32	1	16	720919	40	2	21	730919		27	2	16
710920	51	3	32	720920	47	1	24	730920		43	1	25
710921	40	1	31	720921	43	3	26	730921		56	2	30
710922	45	2	29	720922	39	0	24	730922		47	3	23

as might be expected; but the generally higher number of deaths during the episode days of 1973 are apparent in these two disease categories, as well.

To determine whether morbidity measures parallel the peaking effect noted in the mortality data for the episode period, retrievals were accomplished from the 1971-1973 Medicare files which give a variety of details about hospitalizations during the time in question in Allegheny County, as well as each of the other counties of the Pittsburgh Standard Metropolitan Statistical Area (Beaver, Washington, and Westmoreland counties).

One retrieval provided data on the number of admissions for each day of the 39-day period from August 13 through September 20--the episode days are the middle 13--for the following respiratory disease categories: acute respiratory infections, except influenza (ICDA 460-466), influenza (470-474), pneumonia (480-486), bronchitis, emphysema, and asthma (490-493), other disease of the respiratory system (500-519), and all respiratory disease (460-519).

Tables IV-20, IV-21, and IV-22 summarize the first retrieval of morbidity for Allegheny County. For the pneumonia and bronchitis/emphysema/asthma disease categories (ICDA 480-486 and 490-493, respectively), the counts of admissions are higher during the episode period than during any of the other (eight) 13-day intervals for which data is displayed. In those retrievals showing the classification for all respiratory diseases (ICDA 460-519), the numbers pertaining to the episode, likewise, appear to be significantly larger.

A second retrieval indicated the number of first admissions for each respiratory disease category during the same 39 days of August and September (i.e., only the first hospitalization of a claimant history in a disease grouping was recorded). This retrieval is reproduced in part in Table IV-23. The findings are consistent. During the episode, total admissions for respiratory diseases were higher, and the total of first admissions alone was higher than the eight other 13-day periods.

For a third retrieval, the 39 days were divided into three 13-day intervals and the number of people admitted for respiratory disease during each of

TABLE IV-20

DAILY ADMISSIONS BY DISEASE CLASS IN ALLEGHENY COUNTY, PA, DURING THE PERIOD AUGUST 13 THROUGH SEPTEMBER 20, 1971

	ICDA Group					
Date	460-466x[1]	470-474x[2]	480-486x[3]	490-493x[4]	500-519x[5]	460-519x[6]
Aug 13	0	0	0	0	0	0
14	0	0	0	0	1	1
15	0	0	0	1	1	2
16	0	0	1	0	1	2
17	0	0	0	0	1	1
18	0	0	1	0	0	1
19	0	0	0	1	1	2
20	0	0	1	0	0	1
21	0	0	0	0	1	1
22	1	0	1	2	0	4
23	0	0	1	1	0	2
24	0	0	0	0	0	0
25	0	0	1	1	0	2
26	0	0	0	0	1	1
27	0	0	0	1	0	1
28	0	0	0	0	0	0
29	0	0	0	2	0	2
30	0	0	0	0	0	0
31	0	0	0	0	0	0
Sept 1	0	1	0	1	1	3
2	0	0	0	1	0	1
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	2	2
6	0	0	1	0	1	2
7	0	0	0	0	1	1
8	0	0	2	0	1	3
9	0	0	1	0	1	2
10	0	0	0	1	1	2
11	0	0	0	2	0	2
12	0	0	0	0	0	0
13	0	0	0	0	0	0
14	0	0	0	2	0	2
15	0	0	0	0	0	0
16	0	0	0	0	0	0
17	1	0	0	0	0	1
18	0	0	1	0	0	1
19	1	0	1	0	1	3
20	0	0	1	0	0	1

[1] Acute respiratory infections, except influenza
[2] Influenza
[3] Pneumonia
[4] Bronchitis, emphysema, and asthma
[5] Other diseases of the respiratory system
[6] All diseases of the respiratory system

TABLE IV-21

DAILY ADMISSIONS BY DISEASE CLASS IN ALLEGHENY COUNTY, PA, DURING THE PERIOD AUGUST 13 THROUGH SEPTEMBER 20, 1972

	ICDA Group					
Date	460-466x[1]	470-474x[2]	480-486x[3]	490-493x[4]	500-519x[5]	460-519x[6]
Aug 13	0	0	1	0	0	1
14	0	0	0	0	0	0
15	0	0	0	0	1	1
16	0	0	1	0	1	2
17	0	0	0	0	0	0
18	1	0	1	1	0	3
19	0	0	0	0	0	0
20	0	0	0	0	0	0
21	0	0	0	0	0	0
22	0	0	1	1	1	3
23	0	0	0	0	0	0
24	0	0	2	0	0	2
25	0	0	1	0	1	2
26	1	0	0	0	0	1
27	1	0	0	0	1	2
28	0	0	0	1	0	1
29	0	0	0	2	2	4
30	0	0	0	1	0	1
31	0	0	1	0	2	3
Sept 1	0	0	1	0	1	2
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	1	0	1	0	0	2
6	0	0	0	2	0	2
7	0	0	2	0	1	3
8	1	0	0	1	1	3
9	0	0	0	0	0	0
10	0	0	1	1	0	2
11	0	0	0	2	2	4
12	0	0	0	1	0	1
13	0	0	1	0	0	1
14	0	0	2	1	1	4
15	0	0	0	1	2	3
16	1	0	0	0	1	2
17	0	0	0	0	0	0
18	0	0	0	0	1	1
19	0	0	0	0	0	0
20	0	0	1	0	0	1

[1] Acute respiratory infections, except influenza

[2] Influenza

[3] Pneumonia

[4] Bronchitis, emphysema, and asthma

[5] Other diseases of the respiratory system

[6] All diseases of the respiratory system

TABLE IV-22

DAILY ADMISSIONS BY DISEASE CLASS IN ALLEGHENY COUNTY, PA, DURING THE PERIOD AUGUST 13 THROUGH SEPTEMBER 20, 1973

	ICDA Group					
Date	460-466x[1]	470-474x[2]	480-486x[3]	490-493x[4]	500-519x[5]	460-519x[6]
Aug 13	0	0	1	0	0	1
14	0	0	1	0	0	1
15	0	0	0	0	1	1
16	1	0	0	0	2	3
17	0	0	3	0	0	3
18	0	0	0	1	1	2
19	0	0	0	0	2	2
20	3	0	0	1	0	4
21	0	0	0	2	3	5
22	0	0	0	0	1	1
23	1	0	0	0	0	1
24	0	0	1	0	2	3
25	1	0	1	0	0	2
26	0	0	0	0	2	2
27	0	0	2	0	1	3
28	0	0	0	0	2	2
29	0	0	0	0	2	2
30	0	0	4	3	1	8
31	0	0	0	1	0	1
Sept 1	0	0	0	0	1	1
2	1	0	1	1	0	3
3	0	0	0	1	0	1
4	0	0	1	2	1	4
5	1	0	2	0	1	4
6	0	0	2	0	0	2
7	0	0	1	0	1	2
8	0	0	1	0	0	1
9	0	0	0	0	1	1
10	0	0	0	0	0	0
11	1	0	1	0	0	2
12	0	0	0	1	1	2
13	0	0	0	0	0	0
14	0	0	0	1	1	2
15	1	1	0	0	0	2
16	1	0	0	0	2	3
17	0	0	0	1	0	1
18	0	0	0	0	1	1
19	0	0	0	0	3	3
20	0	0	0	0	0	0

[1] Acute respiratory infections, except influenza
[2] Influenza
[3] Pneumonia
[4] Bronchitis, emphysema, and asthma
[5] Other diseases of the respiratory system
[6] All diseases of the respiratory system

TABLE IV-23. DAILY COUNTS OF ALLEGHENY COUNTY FIRST ADMISSIONS DURING AUGUST 13-SEPTEMBER 10, 1971, 1972, 1973 BY RESPIRATORY DISEASE CATEGORY (I.E., EACH ADMISSION WAS THE FIRST ADMISSION FOR THE RESPECTIVE DISEASE DURING THE 39-DAY PERIOD FOR THE CLAIMANT)

NOTE: September 11-20 counts are not statistically different and have not been reproduced.

ICDA Category	Aug. 13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	Sept. 1	2	3	4	5	6	7	8	9	10
														1971															
460 466X	00	00	00	00	00	00	00	00	00	01	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
470 474X	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	01	00	00	00	00	00	00	00	00	00
480 486X	00	00	00	01	00	01	00	01	00	01	01	00	01	01	00	00	00	00	00	00	00	00	00	00	01	00	01	01	00
490 493X	00	00	01	00	00	00	01	00	00	02	01	00	01	00	01	00	02	00	00	01	01	00	00	00	00	00	00	00	01
500 519X	00	01	01	01	01	00	01	00	01	00	00	00	00	01	00	00	00	00	00	01	00	00	00	02	01	01	01	01	01
460 519X	00	01	02	02	01	01	02	01	01	04	02	00	02	01	01	00	02	00	00	03	01	00	00	02	02	01	02	02	02
														1972															
460 466X	00	00	00	00	00	01	00	00	00	00	00	00	00	01	01	00	00	00	00	00	00	00	00	01	00	00	01	00	00
470 474X	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
480 486X	01	00	00	01	00	01	00	00	00	01	00	02	01	00	00	00	00	00	01	01	00	00	00	01	00	02	00	00	01
490 493X	00	00	00	00	00	01	00	00	00	01	00	00	00	00	00	01	02	01	00	00	00	00	00	00	02	00	01	00	01
500 519X	00	00	01	01	00	00	00	00	00	01	00	00	01	00	01	00	02	00	02	01	00	00	00	00	00	01	01	00	00
460 519X	01	00	01	02	00	03	00	00	00	03	00	02	02	01	02	01	04	01	03	02	00	00	00	02	02	03	03	00	02
														1973						EPISODE									
460 466X	00	00	00	01	00	00	00	03	00	00	01	00	01	00	00	00	00	00	00	00	01	00	00	01	00	00	00	00	00
470 474X	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00
480 486X	01	01	00	00	03	00	00	00	00	00	00	01	01	00	02	00	00	04	00	00	01	00	01	02	02	01	01	00	00
490 493X	00	00	00	00	00	01	00	01	02	00	00	00	00	00	00	00	00	03	01	00	01	01	02	00	00	00	00	00	00
500 519X	00	00	01	02	00	01	02	00	03	01	00	02	00	02	01	02	02	01	00	01	00	00	01	01	00	01	00	01	00
460 519X	01	01	01	03	03	02	02	04	05	01	01	03	02	02	03	02	02	08	01	01	03	01	04	04	02	02	01	01	00

TABLE IV-24

NUMBERS OF PERSONS ADMITTED WITH INDICATED DIAGNOSES EACH 13-DAY PERIOD: AUGUST 13-25, AUGUST 26-SEPT. 7, AND SEPT. 8-SEPT. 20, 1971-1973

	1971			1972			1973		
Class*	Aug. 13-25	Aug 26-Sept. 7	Sept. 8-20	Aug. 13-25	Aug. 26-Sept. 7	Sept. 8-20	Aug. 13-25	Aug. 26-Sept. 7 (Episode)	Sept. 8-20
140-239	23	44	48	36	42	39	44	44	54
460-466	01	00	02	01	03	02	06	02	03
470-474	00	01	00	00	00	00	00	00	01
480-486	06	01	06	07	05	05	07	13	02
490-493	06	05	05	02	06	07	04	08	03
500-519	06	06	04	04	07	08	12	12	09
520-577	41	44	53	48	44	43	47	36	40

*ICDA Classes

140-239	Neoplasms
460-466	Acute respiratory infections, except influenza
470-474	Influenza
480-486	Pneumonia
490-493	Bronchitis, emphysema, and asthma
500-519	Other respiratory diseases
520-577	Diseases of the digestive system

the 13-day time periods were counted. Table IV-24 shows these counts. During the episode period, 23 persons were hospitalized for respiratory diseases. This was the largest number, by far. The next largest number of respiratory hospitalizations was 17, the 13-day period just preceding the episode. Medicare hospitalizations for respiratory diseases appear related to the quality of air that characterizes an episode.

Finally, for those individuals who were admitted during the 13-day episode interval with a respiratory disease, counts were taken of those who died following a final hospital admission before the end of the same calendar year. That portion of the calendar year was divided into episode period, the remainder of the month of September, and the months of October, November, and December.

During each year 1971 and 1972 there was no Medicare respiratory disease death during the interval corresponding to the 1973 episode, while there were four such deaths during the actual episode in 1973. The only other deaths indicated by this retrieval were one for respiratory disease in November 1971, and one for digestive disease during December 1971. There were no other Medicare deaths.

An interesting example of the connection between morbidity and mortality is found in the four individuals (4) admitted to a hospital for treatment during the 1973 episode who died of a respiratory disease before discharge. Only three claimants admitted for a respiratory illness during the same time span in either 1971 or 1972 died of respiratory disease following hospital admission in the remaining three months of the respective calendar years. However, the small numbers do not justify conclusions.

Admissions for neoplasm and digestive disorders do not show a peaking of admissions during the pollution episode, such as that found for the respiratory disease categories. This finding was expected. It was performed as one of several tests of the credibility of the Medicare morbidity data. These two sets of admissions data are not reproduced or discussed further.

Expanding attention from Allegheny County to the larger geographical area, Table IV-25 presents admissions statistics from all four counties of the Pittsburgh SMSA for pneumonia and bronchitis/emphysema/asthma in each of the years, 1971, 1972, and 1973. While the counts from Beaver, Washington, and Westmoreland counties do not, in the absence of data from Allegheny County, exhibit any pronounced peaking during the episode days of 1973, the daily total for residents for the SMSA (Allegheny included) does not obscure the elevated levels of admissions for the respiratory ailments during that time interval. The total number of admissions for pneumonia during the episode is about twice the average of the other eight intervals; for bronchitis/emphysema/asthma it is also about twice the average for the other eight intervals.

An indeterminate number of the higher quantity of admissions may be attributed to the unusually warm Allegheny County temperatures that coincided with the 13-day episode. Episode days numbers 1, 12, and 13 had maximum temperatures of less than 90° Fahrenheit; days 2 through 11 experienced 90° and above. By contrast, in only two days of the 104 days comprising the other eight intervals were there maximum temperatures of 90° or above. Data for northeastern and midwestern U.S. cities indicate higher mortalities of all kinds occur as temperatures exceed 80°, and it is reasonable to include the coincidental higher temperature as a factor contributing to the increased respiratory disease hospitalizations noted in Table IV-25.

Summarizing this section, various Medicare morbidity measures did reflect a peaking effect during the Pittsburgh air pollution episode in the same way that mortality data did. As indicated in several tabulations, total admissions for pneumonia, bronchitis/emphysema/asthma, and all respiratory disease increased during the episode, as did counts of first admissions for these respiratory diseases. There is also some indication of a "build-up" in respiratory disease hospitalizations coincident with the time period immediately preceding the air pollution episode. With a 20 percent sample, the three categories of respiratory disease measures are essentially the same. In a larger sample, small differences would be expected, but each should reflect similar health effects.

TABLE IV-25

DAILY ADMISSIONS FOR PNEUMONIA AND BRONCHITIS, EMPHYSEMA, OR ASTHMA IN THE PITTSBURGH SMSA DURING THE 39-DAY PERIOD INCLUDING THE 1973 AIR POLLUTION EPISODE AND COMPARABLE PERIOD IN 1971 AND 1972

	Pneumonia				Bronchitis, Emphysema, or Asthma			
	1971	1972	1973		1971	1972	1973	
Aug. 13	1	1	1		0	1	0	
14	0	1	1		0	0	0	
15	0	0	0		1	0	0	
16	2	1	0		1	0	0	
17	1	0	3		0	0	0	
18	2	1	0		0	1	2	
19	1	0	0		1	0	0	
20	1	0	0		0	0	1	
21	0	0	0		0	1	2	
22	1	1	0		3	1	0	
23	2	0	1		2	1	0	
24	0	3	1		1	0	0	
25	1	1	2		1	0	0	
TOTAL	12	9	9		10	5	5	
Aug. 26	0	0	0		0	0	1	
27	0	0	3		1	0	1	
28	0	1	0		0	1	0	
29	0	0	0		2	2	2	
30	1	0	4		0	1	3	
31	0	1	0		2	0	2	
Sept. 1	0	1	3	Episode	1	0	1	Episode
2	0	1	1		1	0	1	
3	0	0	0		0	0	1	
4	0	0	1		0	0	2	
5	0	2	2		0	0	0	
6	2	0	2		0	2	0	
7	0	2	2		0	0	0	
TOTAL	3	8	18		7	6	14	
Sept. 8	2	0	2		0	1	0	
9	2	0	1		1	0	1	
10	0	1	0		1	1	0	
11	1	0	3		2	2	1	
12	1	0	0		0	1	1	
13	0	1	0		0	0	0	
14	0	2	0		2	1	1	
15	0	0	0		0	1	1	
16	0	0	0		0	0	0	
17	0	0	0		0	0	1	
18	1	2	1		1	0	1	
19	1	1	2		0	0	0	
20	1	1	1		1	0	0	
TOTAL	9	8	10		8	7	7	

The Pittsburgh data indicate that increased respiratory disease hospitalizations can be expected to occur during an air pollution episode. On the other hand, analyses of Medicare hospitalizations for digestive and neoplasm diseases showed no relationships to air pollution episode timing. These two findings indicate the Medicare files have sufficient reliability as indicators of the effects of the environment on human health to justify further investigation.

c. "Survival" Times as an Indication of Medicare Data Reliability--

The Medicare hospital discharge survey file is not the only source of national morbidity statistics, although none can match Medicare's extensiveness and detail. Bi-weekly reports on respiratory and other diseases are published by the Center for Disease Control, and cancer incidence data are compiled by the National Cancer Institute (NCI) from information supplied by tumor registries in nine selected states across the United States. Both are widely used. The NCI registry attempts to be particularly thorough for the areas in which it obtains data.

It is feasible to assess the representativeness and utility of the Medicare data for a variety of epidemiological research purposes by comparing findings from these other sources with data on the same subject contained in the 20 percent 1971-1973 Medicare file sample. Two comparisons were made. First, average "survival" times from first cancer diagnosis in a claimant's hospitalization history to the time of death are compared with findings of the NCI Surveillance, Epidemiology and End Results Reportings (SEER) program. Second, the Medicare and NCI data were ranked in terms of deaths related to total patients with a specific neoplasm.

It is noted that NCI uses a somewhat different coding procedure and its data include cases that pre-date the Eighth ICDA coding used in the Medicare data. To reduce possibility of errors, ten NCI cancer types were selected that are believed most directly comparable, and least likely to be involved with coding complexities. NCI classifies the population of Medicare age under two age classes; 65 to 74 years, and over 75 years of age. These differences are taken into consideration, as discussed below.

A comparative "survival time" methodology identical with that presented herein may be applied with respiratory diseases in a context of determining the intervals between first admission and death for parts of the country that differ in environmental exposures. The methodology may also have applications in evaluating the comparative severity of air pollution episodes in different parts of the country, and/or the severity of a particular episode in reducing the time-to-death interval below "normal" for that area.

Time-to-death--A retrieval from the Medicare sample data file was designed to provide information on the numbers of individuals with a cancer diagnosis who die and the approximate time span between the first admission (during 1971-1973) for a diagnosed cancer and the last admission, following which the claimant is discharged dead due to either a cancer or non-cancer cause.

This retrieval is illustrated in part by Table IV-26, which reproduces most of the detail of the first page of the printout. As shown in the examples in Table IV-26, a distinction is made between first admissions for cancer which are the same as the specified ICDA-code cause of death and those admissions for which the discharge diagnoses are in a different cancer category than that ultimately causing death. A third line of data shows the distribution of times from the first admission for the cancer to which death is attributed to the time of the claimant's last admission (that is, cases in which an admission for a different cancer category occurred earlier in the claimant's hospitalizations history). The fourth line of data shows cancer-diagnosed claimants dying of non-cancer causes.

These details are significant for epidemiological studies, and some are particularly relevant to possible relationships between the environment and condition of human health. For example, environmental factors have been variously evaluated as causing 60 to 90 percent of cancer, and this Medicare morbidity/mortality file offers the only source for including that proportion of cancer cases which terminate in death from non-cancer causes. It is also noted from the four examples in Table IV-26, that considerable numbers of

TABLE IV-26

ILLUSTRATIVE DISTRIBUTIONS OF TIME SPANS FROM FIRST ADMISSION FOR CANCER TO LAST ADMISSION (DISCHARGED DEAD), FROM CONSOLIDATED MEDICARE RECORDS, 1971-1973

ICDA	ALL ADMISSIONS	NO. DIED	ADMITTED EARLIER OTHER CANCER	TIME SPAN (MONTHS) 7-9	10-12	13-18	19-24	Total Cases*
ICD140-145	0005903	0000461	0000119					
TIME FROM 1ST ADMISSION-OTHER CANCER				0009	0011	0013	0010	94
TIME FROM 1ST ADMISSION-SAME CANCER				0029	0022	0025	0010	100
TIME FROM 1ST ADMISSION-SAME WHEN OTHER PRECEDES				0005	0002	0000	0001	23
TIME FROM 1ST ADMISSION-FOR CANCER: NON-CANCER DEATH				0007	0006	0012	0003	61

ICDA	ALL ADMISSIONS	NO. DIED	ADMITTED EARLIER OTHER CANCER	TIME SPAN (MONTHS) 7-9	10-12	13-18	19-24	
ICD146-149	0002020	0000291	0000105					
TIME FROM 1ST ADMISSION-OTHER CANCER				0016	0012	0008	0005	95
TIME FROM 1ST ADMISSION-SAME CANCER				0016	0011	0003	0002	84
TIME FROM 1ST ADMISSION-SAME WHEN OTHER PRECEDES				0002	0000	0001	0002	22
TIME FROM 1ST ADMISSION-FOR CANCER: NON-CANCER DEATH				0003	0000	0003	0002	15

ICDA	ALL ADMISSIONS	NO. DIED	ADMITTED EARLIER OTHER CANCER	TIME SPAN (MONTHS) 7-9	10-12	13-18	19-24	
ICD151-151	0010213	0002601	0000323					
TIME FROM 1ST ADMISSION-OTHER CANCER				0029	0016	0022	0009	263
TIME FROM 1ST ADMISSION-SAME CANCER				0138	0069	0074	0031	1004
TIME FROM 1ST ADMISSION-SAME WHEN OTHER PRECEDES				0007	0004	0001	0000	60
TIME FROM 1ST ADMISSION-FOR CANCER: NON-CANCER DEATH				0016	0007	0011	0003	91

ICDA	ALL ADMISSIONS	NO. DIED	ADMITTED EARLIER OTHER CANCER	TIME SPAN (MONTHS) 7-9	10-12	13-18	19-24	
ICD153-153	0033602	0006120	0001081					
TIME FROM 1ST ADMISSION-OTHER CANCER				0108	0096	0113	0065	874
TIME FROM 1ST ADMISSION-SAME CANCER				0267	0170	0210	0103	1869
TIME FROM 1ST ADMISSION-SAME WHEN OTHER PRECEDES				0024	0018	0011	0006	229
TIME FROM 1ST ADMISSION-FOR CANCER: NON-CANCER DEATH				0041	0024	0039	0022	263

* Time spans for less than seven months, and more than 24 months, have been deleted from this illustrative printout. The time spans shown will not sum to the total cases column. The time span column 0-3 months includes first admission cases discharged dead.

cancer hospitalizations are for cancer types different from the cancer type that eventually terminates in a cancer death.

This Medicare information retrieval provided the basis for Table IV-27 which shows the average time span in months for the ten selected NCI-Medicare disease categories from first Medicare cancer hospitalization record to death. Also shown is the relative rank in parentheses of the particular cause of death. It is noted that the Medicare tapes available were for only three years, 1971-1973, and some cancers (prostate, bladder, colon, and rectum) have comparatively long median survival times. Patients may therefore have experienced Medicare hospitalizations prior to 1971, but for purposes of this retrieval it was assumed that the earliest cancer admission listed on the tape was their first cancer hospitalization admission. The net result of this assumption is to understate the average time span from first admission to death. For cancer types in which death occurs relatively quickly, the understatement is not significant. However, for those cancers with average or median survival times of several years, an average survival time calculated from tapes showing 36 months at most would necessarily underestimate the actual "survival" duration. Nevertheless, if this qualification is kept in mind, and supplemented with ranking of the time intervals from "first" admission to death, the findings should provide a basis for evaluating consistency of Medicare files with NCI files for the same cancer types.

Table IV-27 presents, by cancer disease category, the average number of months from actual or assumed first Medicare cancer admission to death. Separate averages were calculated where more than one cancer type was involved. For those cases where death is known to have occurred, although for a non-cancer cause, the average interval to death was also calculated. Rankings are shown within parentheses for each of the four broad "first-admission-to-death" classes.

The Medicare rankings as calculated and shown in Table IV-27 are almost identical with the NCI rankings of median survival times for these same diseases. There is also a close consistency between the median survival times for each cancer type, as calculated for NCI's experience, and the average

TABLE IV-27

AVERAGE MONTHS AND RANK FROM FIRST MEDICARE CANCER TO DEATH OF VARIOUS CAUSES

Disease Category	(a) Months Rank	(b) Months Rank	(c) Months Rank	(d) Months Rank
Pharynx 146-149*	7.22 (5)	5.83 (6)	4.86 (6)	10.93 (9)
Stomach 151*	6.52 (2)	5.78 (5)	4.18 (3)	7.18 (4)
Large Intestine (Exc. rectum) 153*	8.67 (7)	7.47 (7)	5.19 (7)	8.88 (6)
Rectum/Recto-sigmoid Junction 154*	8.36 (6)	8.19 (8)	7.12 (9)	9.66 (7)
Pancreas 157*	5.08 (1)	5.62 (3)	2.66 (1)	6.42 (1)
Trachea/Bronchus/Lung 162*	7.14 (4)	5.14 (2)	4.32 (5)	6.58 (2)
Prostate 185*	8.76 (8)	9.34 (10)	6.18 (8)	9.85 (8)
Bladder 188*	10.11 (10)	9.23 (9)	7.96 (10)	11.37 (10)
Brain 191*	6.64 (3)	4.00 (1)	3.64 (2)	8.00 (5)
Leukemia 204-207*	10.06 (9)	5.74 (4)	4.24 (4)	7.15 (3)

*Causes of Death Selected for Comparability with NCI "End Result In Cancer."

(a) First cancer admission is other than cancer cause of death

(b) First cancer admission is same as cause of death

(c) Time from first admission for same cancer as cause of death when first mention is for other cancer

(d) Non-Cancer Death

Categories of first admission for cancer given in left margin survival times derived from the Medicare tapes for these same types of cancer. Although the Medicare and NCI ranks are essentially the same where the first cancer diagnosis/admission is also the cause of death, the average months to death calculated from the Medicare tapes are significantly less than the NCI median number of months for four longer survival types of cancer--prostate, bladder, colon, and rectum. As indicated above, the 1971-1973 Medicare tape limits cause this difference to be expected.

NCI/Medicare rankings of "survival" times--Table IV-28 presents the median survival times for the ten cancer categories as calculated from NCI files for NCI patients 65 years of age and over. The median survival times shown in column (e) can be related directly to column (b) of Table IV-27, as both are concerned with the survival time from diagnosis/admission of the cancer type that eventually caused death. The times are comparable with the limitations of the Medicare data taken into account, as discussed above.

Table IV-28 also presents relative rankings, in parentheses, from NCI and Medicare patient descriptors files, by cancer type. The NCI rankings are of median survival times; the Medicare rankings are the highest to lowest percentages of deaths among those admitted with indicated cancer in Medicare records. These percentages of death-to-admissions varied from 47.5 percent (ICD 157, pancreas) to 13.0 percent (ICD 188, bladder). The percentages were calculated on the assumption that the higher the percentage of deaths for particular cancer types, the lower the survival time from hospitalization diagnosis to death. As Table IV-28 shows, this assumption is almost perfectly correlated with the NCI rankings of median survival times.

In summary, there is consistency, and a close relationship between Medicare morbidity/mortality data and comparable tabulations of the National Cancer Institute for those categories of cancer defined in identical or near-identical terms. Medicare provides a nationwide sample, and NCI offers more thorough coverage in relatively restricted areas. Neither Medicare nor NCI purport to have complete files. Available data are not definitive as to which is more "correct." Medicare morbidity/mortality files provide data not available through NCI that are relevant to EPA's environmental health mission, and

TABLE IV-28

CALCULATION OF MEDIAN SURVIVAL TIMES FOR CANCER PATIENTS AGE 65 AND OLDER, AND COMPARISON OF RELATIVE NCI AND MEDICARE RANKINGS

NCI Neoplasm Subclasses	Assumed ICDA #	Number of Patients		Survival Time (yrs)		Medical Survival Time (yrs)	Relative Rank	
		65-74 (a)	over 75 (b)	65-74 (c)	over 75 (d)	(e)	NCI (f)	Medicare (g)
Pharynx	146-149	801	436	1.0	0.9	1.0	(6)	(7)
Stomach	151	3,490	2,722	0.4	0.3	0.4	(5)	(4)
Colon	153	6,324	5,215	2.2	1.0	1.7	(8)	(6)
Rectum	154	3,760	2,878	1.9	1.0	1.5	(7)	(8)
Pancreas	157	1,768	1,292	0.2	0.2	0.2	(1)	(1)
Lung and Bronchus	162	6,783	2,304	0.4	0.2	0.3	(4)	(2)
Prostate	185	5,382	5,694	3.6	2.0	2.8	(10)	(9)
Bladder	188	3,474	2,916	3.0	1.4	2.3	(9)	(10)
Brain	191	1,990$^+$		0.3$^+$		0.3	(3)	(5)
Leukemia	204-207	459	955	0.2	0.2	0.2	(1)	(3)

(a) Number of patients age 65-74 diagnosed with disease, 1955-1964.*

(b) Number of patients age 75 and older diagnosed with disease, 1955-1964.*

(c) Median survival time (years) for patients age 65-74 diagnosed with disease, 1955-1964.*

(d) Median survival time (years) for patients age 75 and older diagnosed with disease, 1955-1964.*

(e) Median survival time (years) for patients age 65 and older diagnosed with disease, computed from columns (a)-(d).

(f) Relative rank of median survival times in column (e).

(g) Relative rank of percentages of deaths among those admitted with indicated cancer in Medicare records.

+ Reported only for patients age 65 and older.

*From data compiled by End Results Section, Biometry Branch, National Cancer Institute, End Results in Cancer, Report No. 4, DHEW Publication No. NIH 73-272.

it is believed plausible, prudent, and appropriate for EPA to evaluate Medicare tapes generally as possessing at least the same credibility and authoritativeness as NCI tape sources.

6. Medicare Claimant Migration

Linkages between environmental pollution and adverse human health effects are difficult to determine because of the long time lag between exposure and appearance of some effects, and due to movement of individuals in and out of the exposure area during that time interval. Comprehensive, detailed migration data needed to address such moves explicitly do not now exist in adequate detail, although steps are underway to remedy this deficiency in significant respects. The Medicare hospital discharge file, showing county of residence at the time of each admission, as well as age, ICDA code, and other factors, may be of use in determining migration patterns of a population believed likely to exhibit the delayed effects of environmental exposures.

To provide some indication of the migration analysis value of the file, retrievals were made from the merged Medicare files (which had been sorted by claim identification code) to determine relocations of residents who, at the time of at least one Medicare admission for treatment during 1971-1973, were resident in one of the seven study counties in Florida and Pennsylvania. Counts were made of the number of counties and number of states listed as place of residence for each claimant over the three-year period. In addition, the number of counties within Florida or Pennsylvania given as place of residence were shown. Enumerations were performed separately for males and females, with a further subdivision for white or non-white race. Tables were prepared by county of residence, as well as by race-sex category. Table IV-29 summarizes these statistics.

Clearly, a vast majority of the claimants in this three-year data sample show a single county of residence. Only the non-white residents of Pinellas County, Florida, have a migration rate greater than five percent, and that is based on a very small number of cases. The rates in the Florida counties are

TABLE IV-29

FREQUENCY DISTRIBUTION OF NUMBERS OF (A) COUNTIES OF RESIDENCE, (B) STATES OF RESIDENCE, AND (C) COUNTIES OF RESIDENCE WITHIN INDICATED STATE DURING 1971-1973 FOR MEDICARE CLAIMANTS RESIDING IN INDICATED STATE AND COUNTY

Broward County, Florida

			White		Nonwhite	
			Males	Females	Males	Females
A.	Number	1	2,326	2,043	63	46
	of	2	118	67	2	1
	Counties	3	1	2	0	0
B.	Number	1	2,350	2,066	63	47
	of	2	94	44	2	0
	States	3	1	2	0	0
C.	Number	1	2,421	2,089	65	46
	of	2	24	23	0	1
	Counties	3	0	0	0	0
	within Florida					

Pinellas County, Florida

			White		Nonwhite	
			Males	Females	Males	Females
A.	Number	1	2,606	2,443	28	35
	of	2	90	59	2	2
	Counties	3	2	3	0	0
B.	Number	1	2,632	2,452	28	35
	of	2	75	52	2	2
	States	3	1	1	0	0
C.	Number	1	2,691	2,494	30	37
	of	2	17	11	0	0
	Counties	3	0	0	0	0
	within Florida					

Berks County, Pennsylvania

			White		Nonwhite	
			Males	Females	Males	Females
A.	Number	1	569	545	6	5
	of	2	2	9	0	0
	Counties	3	0	0	0	0
B.	Number	1	570	548	6	5
	of	2	1	6	0	0
	States	3	0	0	0	0
C.	Number	1	570	551	6	5
	of	2	1	3	0	0
	Counties	3	0	0	0	0
	in Pennsylvania					

TABLE IV-29 (continued)

Lackawanna County, Pennsylvania

			White		Nonwhite	
			Males	Females	Males	Females
A.	Number	1	520	513	3	0
	of	2	5	0	0	0
	Counties	3	0	0	0	0
B.	Number	1	523	513	3	0
	of	2	2	0	0	0
	States	3	0	0	0	0
C.	Number	1	522	513	3	0
	of	2	3	0	0	0
	Counties in Pennsylvania	3	0	0	0	0

Lancaster County, Pennsylvania

			White		Nonwhite	
			Males	Females	Males	Females
A.	Number	1	540	544	4	6
	of	2	10	5	0	0
	Counties	3	0	0	0	0
B.	Number	1	545	556	4	6
	of	2	5	3	0	0
	States	3	0	0	0	0
C.	Number	1	545	557	4	6
	of	2	5	2	0	0
	Counties in Pennyslvania	3	0	0	0	0

Luzerne County, Pennsylvania

			White		Nonwhite	
			Males	Females	Males	Females
A.	Number	1	830	811	1	0
	of	2	13	14	0	0
	Counties	3	0	0	0	0
B.	Number	1	831	815	1	0
	of	2	12	10	0	0
	States	3	0	0	0	0
C.	Number	1	842	821	1	0
	of	2	1	4	0	0
	Counties in Pennsylvania	3	0	0	0	0

TABLE IV-29 (continued)

Schuylkill County, Pennsylvania

		White		Nonwhite	
		Males	Females	Males	Females
A. Number of Counties	1	459	432	0	1
	2	5	1	0	0
	3	0	0	0	0
B. Number of States	1	461	432	0	1
	2	3	1	0	0
	3	0	0	0	0
C. Number of Counties in Pennsylvania	1	462	433	0	1
	2	2	0	0	0
	3	0	0	0	0

generally greater than those in the Pennsylvania counties. No hospitalization history gives more than three residence addresses. The number of individuals admitted to hospitals in more than one study county is inconsequential.

The examination of Medicare claimant migration did not indicate as widespread a relocation pattern as one might expect from the retirement age population, especially in the high in-migration counties of Florida. It may be that most individuals who change residences do so shortly after the retirement date, while still in relatively good health. The data do indicate that once hospitalization occurs, few individuals change their residence.

The Medicare records are, therefore, not a good source of data either on migration patterns, or on the contribution of previously incurred environmental/occupational exposures to health statistics of various geographic regions. A more complete representation could be found by examining the claimant's employment history by industry and county in conjunction with the Medicare data. Steps have been initiated to obtain this industry/county data for the working population prior to their retirement years, and for linking county/industry of employment with county or counties of retirement.

7. Demonstration--Morbidity/Mortality Analyses in Steel Industry Counties

Considerable interest was expressed within EPA for determining whether the techniques and information being used in this project could be applied to an investigation of a single industry on a county level of detail. The steel industry was of particular concern because of litigation initiated by steel industry representatives concerning EPA enforcement of a standard for Total Suspended Particulate. Under the direction of the EPA Project Officer, the morbidity experience of the Medicare population in steel-producing counties was featured in the example application. It is emphasized that the work on this task is designed only to demonstrate the versatility of the data bases for EPA purposes and not to establish the validity of arguments on either side of the issue under contention.

a. Identification of Steel Industry Counties and Data Collection--

The work involved identification of the major steel-producing counties of the United States, and comparisons of demographics, pollutant emissions, air quality, morbidity and mortality rates with U.S. totals and statistics from steel-producing and non-steel-producing counties. A list of locations of steel-producing establishments meeting the requirements for Standard Industrial Classification 3312 was provided by EPA. This classification includes blast furnaces (including coke ovens), steel works, and rolling mills. A compilation of the counties with establishments classified under SIC 3312 is given in Table IV-30; there were 97 such counties in 31 states. Figure IV-7 shows the geographic distribution of these counties (except Honolulu, Hawaii).

The form used for entry of all relevant data elements is shown in Table IV-31. The first of its six sections is administrative, containing the state and county names, the EPA/SAROAD number (the identifier used by the SSI-developed POPATRISK data base), and the Social Security Administration state-county number used with its Medicare hospitalization survey files.

The second section contains Bureau of Census county demographic information retrieved from POPATRISK: total county population in 1970, county population divided by county land area, estimated percent of population living within town or city boundaries, percent of population that is white, percent female, percent over the age of 64 years, absolute number of white males and white females over the age 64, and proportion of white males and white females over 64 years of age in 1970 who resided (a) outside the county in 1965 and within the county in 1970, or (b) inside the county in 1965 and outside the county in 1970.

The third section contains totals for emissions (in tons per year) of TSP and SO_2, the summation of emissions by individual point sources within the county, and the arithmetic means (in $\mu g/m^3$) of the site-specific geometric means of ambient TSP and SO_2 measurements taken over one-hour or twenty-four-hour sampling intervals in 1973 at county monitoring sites designated for population-oriented, source-oriented, and background surveillance, respectively. Emissions values are taken from EPA/NEDS files, monitoring data are from SAROAD.

TABLE IV-30

LIST OF 31 STATES, 97 COUNTIES, TWO INDEPENDENT CITIES OF THE UNITED STATES IN WHICH ARE LOCATED MANUFACTURING ESTABLISHMENTS WITH SIC CODE 3312--BLAST FURNACES (INCLUDING COKE OVENS), STEEL WORKS, AND ROLLING MILLS

(Source: NEDS)

- Alabama
  - Etowah
  - Jefferson
- California
  - Alameda
  - Los Angeles
  - San Bernadino
- Colorado
  - Pueblo
- Connecticut
  - Fairfield
- Delaware
  - New Castle
- Florida
  - Hillsborough
  - Martin
- Georgia
  - Fulton
- Hawaii
  - Honolulu
- Illinois
  - Cook
  - Madison
  - Whiteside
- Indiana
  - Allen
  - Henry
  - Howard
  - Lake
  - Marion
  - Porter
  - Vigo
- Kentucky
  - Boyd
  - Campbell
  - Daviess
- Louisiana
  - Tangipahos Parish
- Maryland
  - Baltimore City
  - Baltimore
- Michigan
  - Macomb
  - Wayne
- Minnesota
  - Ramsey
  - St. Louis
- Missouri
  - St. Louis City
  - St. Louis
- New Jersey
  - Middlesex
- New York
  - Albany
  - Chautauqua
  - Erie
- North Carolina
  - Mecklenberg
- Ohio
  - Ashtabula
  - Butler
  - Cuyahoga
  - Jefferson
  - Lake
  - Lawrence
  - Lorain
  - Lucas
  - Mahoning
  - Richland
  - Scioto
  - Stark
  - Trumbull
  - Washington
- Oklahoma
  - Tulsa
- Oregon
  - Multnomah
  - Yamhill
- Pennsylvania
  - Allegheny
  - Beaver
  - Berks
  - Bucks
  - Butler
  - Cambria
  - Chester
- Pennsylvania (cont.)
  - Dauphin
  - Erie
  - Lawrence
  - Lebanon
  - Mercer
  - Mifflin
  - Montgomery
  - Northampton
  - Northumberland
  - Philadelphia
  - Venango
  - Warren
  - Washington
  - Westmoreland
- South Carolina
  - Darlington
  - Georgetown
  - Richland
- Tennessee
  - Hamilton
  - Knox
  - Roane
- Texas
  - Austin
  - Chambers
  - El Paso
  - Gray
  - Guadalupe
  - Harris
  - Morris
  - Tarrant
- Utah
  - Utah
- Virginia
  - Buchanan
- Washington
  - King
- West Virginia
  - Brooke
  - Cabell
  - Hancock
  - Marion
- Wisconsin
  - Milwaukee

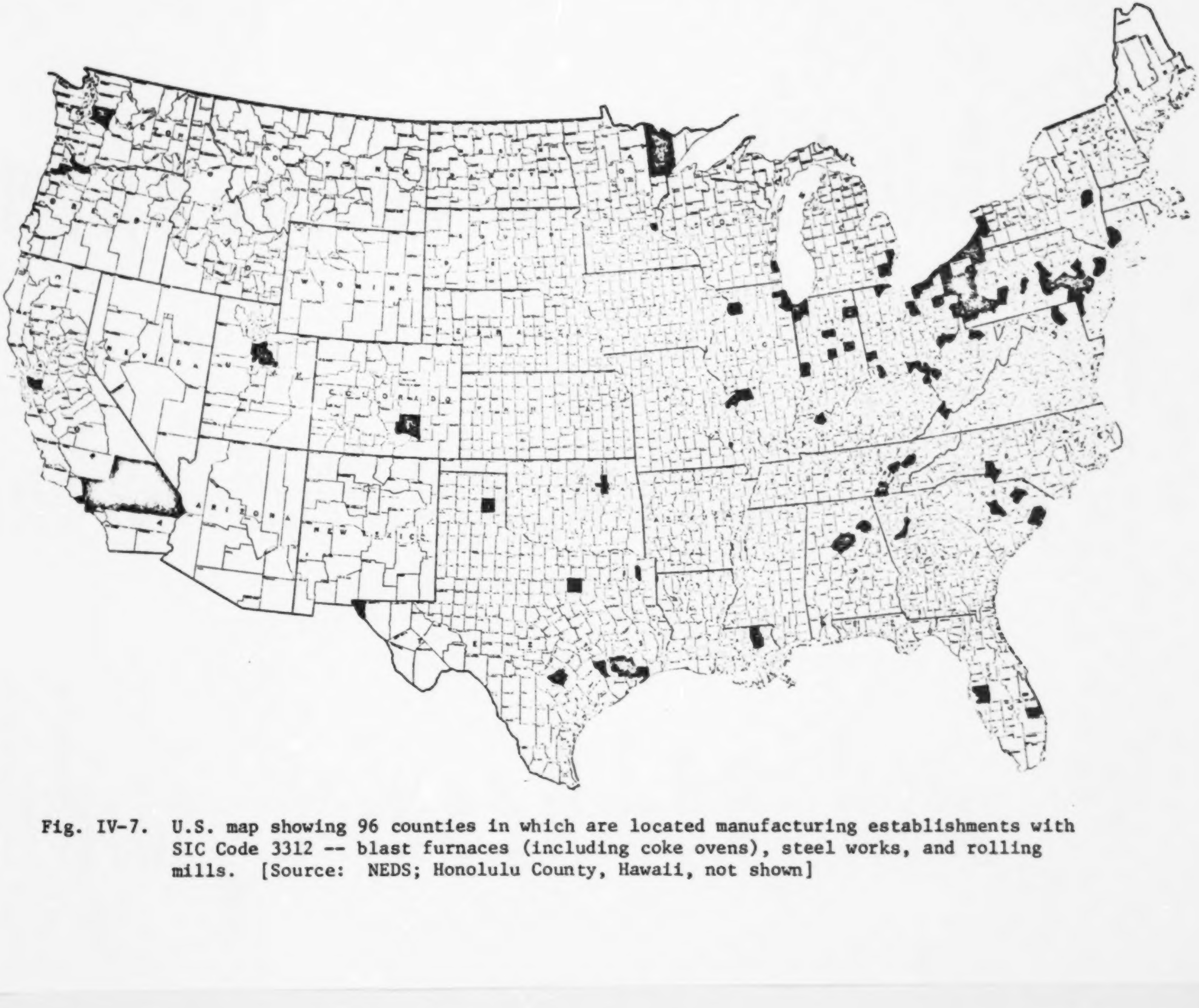

COLORADO
ARIZONA
NEVADA
WYOMING

Fig. IV-7. U.S. map showing 96 counties in which are located manufacturing establishments with SIC Code 3312 -- blast furnaces (including coke ovens), steel works, and rolling mills. [Source: NEDS; Honolulu County, Hawaii, not shown]

TABLE IV-31
DATA ENTRY FORM

Serial # ________

Section 1

________ / ________ / ________ / ________ /
State / County / SAROAD # / SSA #

Section 2

________ / ________ / ________ / ________ / ________ / ________ /
Population / Ppn./Mi.^{2} / % Urban / % White / % Female / % Over 64

________ / ________ / ________ / ________ / ________ / ________ /
Males Over 64 / Females Over 64 / % Male 65+ IN-MIGR / % Female 65+ IN-MIGR / % Male 65+ OUT-MIGR / % Female 65+ OUT-MIGR

Emissions: TSP County Total ______ Point Sources ______
SO_2 County Total ______ Point Sources ______

Monitoring: (Site Means) TSP Population ______ Source ______ Background ______
SO_2 Population ______ Source ______ Background ______

$\frac{TSP}{Mi^2}$ ______ $\frac{SO_x}{Mi^2}$ ______ Sec-tion

% 3312 TSP ______
% 3312 SO_x ______

Section 4

Rank: (In Top 500)
________ / ________ / ________ / ________ /
TSP Emissions / SO_2 Emissions / Mean TSP Predict / Monitoring

Mortality Rates from	White Male	White Female
All Neoplasms (ICDA 140-239)		
Cancer of Buccal Cavity/Pharynx (140-149)		
Cancer of Digestive Organs/Peritoneum (150-159)		
Cancer of Respiratory System (160-163)		
Cancer of Bladder (188)		
Diseases of the Respiratory System (460-519)		
Acute Interstitial/Bronchopneumonia (484,485)		
Bronchitis, Emphysema, Asthma (490-493)		

Section 5

Medicare Morbidity:

	White Claimants Male	White Claimants Female	/ Rates per 1000 Male	Rates per 1000 Female
All Respiratory Diseases (ICDA 460-519)			/	
Acute Infection, Except Influenza (460-466)			/	
Influenza (470-474)			/	
Pneumonia (480-486)			/	
Bronchitis/Emphysema/Asthma (490-493)			/	
Other Respiratory Diseases (500-519)			/	

Section 6

Mortality Rates:	Male	Female
Influenza and Pneumonia (470-486)		
Bronchitis (490-492)		
Asthma (493)		
Other Bronchopulmonic Disease (460-466, 500-519)		
Cancer of Buccal Cavity/Pharynx (140-149)		
Cancer of Digestive Organs/Peritoneum (150-159)		
Cancer of Respiratory System (160-163)		
Cancer of Urinary Organs (188,189)		

Indication is given in the fourth section of each steel-producing county's rank, if applicable, in the top 500 counties of the United States with regard to TSP emissions, SO_2 emissions, mean predicted for TSP monitors (calculated by an algorithm for counties in which TSP and NO_x emissions fall within a specified range), the mean of county population-oriented TSP monitoring sites, and white male/white female mortality rates for all neoplasms, certain subgroupings of malignant neoplasms, and three respiratory disease categories. The emissions and ambient ranks were derived from the POPATRISK data base. The mortality ranks were done under an earlier SSI-EPA contract, "Investigation into the Industrial Correlates of Environment Related Mortality."

Section five presents a compilation by SSI of data concerning white males and white females from the Medicare sample. Statistics were obtained of the number of admissions/discharges and of the number of unique claim numbers with diagnoses of six respiratory disease classifications. The sixth division of the data entry form displays mortality rates, averaged over the years 1969 to 1971 (POPATRISK data base) for white males and white females of all age groups in four respiratory disease and four cancer categories.

b. Comparison of Morbidity in Steel-Producing Counties with State and U.S. Totals--

Total numbers of white male and white female Medicare claimants requiring hospitalization in 1971-1973 for each of the six respiratory illness classifications for the 97 "steel" counties were obtained and divided by the respective total white male and female populations over the age of 64 years. Rates per 1,000 for each of the six disease categories (male and female) were computed and compared with total U.S. ratios and those of the complementary set of U.S. counties (i.e., without data from the 97 "steel" counties). As seen in Table IV-32, for each of the six disease groups for each sex, the steel county morbidity rates are lower than the corresponding total U.S. rates.

These findings should not be construed as evidence that steel industry emissions are not harmful. Individuals more susceptible to respiratory diseases caused or aggravated by a co-located steel plant may have died before

TABLE IV-32

TOTAL NUMBER OF MEDICARE CLAIMANTS (WITH CORRESPONDING MORBIDITY RATES) IN 97 STEEL-PRODUCING COUNTIES AND ALL U.S. COUNTIES FOR HOSPITALIZATIONS IN 1971-1973 WITH DIAGNOSES IN SIX RESPIRATORY DISEASE CATEGORIES

	STEEL COUNTY TOTALS*†		RATES/1000	
	White Males Over Age 64	White Females Over Age 64	White Males 65+	White Females 65+
All Respiratory Diseases (ICDA 460-519)	21,896	19,538	69.1	41.5
Acute Infection, Except Influenza (460-466)	1,907	2,852	6.0	6.1
Influenza (470-474)	721	1,360	2.3	2.9
Pneumonia (480-486)	8,284	7,684	26.1	16.3
Bronchitis/Emphysema/Asthma (490-493)	6,203	4,296	19.6	9.1
Other Respiratory Diseases (500-519)	7,349	4,926	23.2	10.5
Population	1,584,893	2,356,676		

	U.S. TOTALS		RATES/1000	
	White Males Over Age 64	White Females Over Age 64	White Males 65+	White Females 65+
All Respiratory Diseases (ICDA 460-519)	125,224	112,574	81.9	52.7
Acute Infection, Except Influenza (460-466)	14,662	19,910	9.6	9.3
Influenza (470-474)	7,659	12,476	5.0	5.8
Pneumonia (480-486)	47,241	43,589	30.9	20.4
Bronchitis/Emphysema/Asthma (490-493)	36,313	22,651	23.7	10.6
Other Respiratory Diseases (500-519)	36,667	24,808	24.0	11.6
Population	7,645,675	10,684,667		

---

* Excludes counts for Balitmore City and St. Louis City.

† Individual claimants may be counted in more than one county and/or disease classification.

attaining Medicare age/eligibility, or may have relocated to a less hazardous environment. Available resources did not permit introduction of these and other factors affecting conclusions on possible environmental health relationships.

Steel county morbidity rates were then compared with the corresponding rates of their respective states. It is noted that the state rate included data from its steel-producing counties. Table IV-33 indicates the counties with respiratory morbidity rates higher than state rates. It is also noted that for only male disease categories 1 and 4 (i.e., all respiratory disease and pneumonia) was there a majority (actually 51.5 percent) of steel counties with rates greater than their corresponding state rates.

c. Investigation of Relationship Between Morbidity and Pollution Levels--

To determine whether a relationship could be found between manifested health effects and proportion of TSP and SO_x point source emissions attributable to steel production facilities, a retrieval from NEDS keyed to sources with SIC 3312 was accomplished. The emissions values of the steel-producing sources were totaled by county and divided by respective county point source totals. This percentage was then entered on the data entry form, along with computations of TSP and SO_x estimated total emissions per square mile of county geographic area The percentage provided an estimate of the relative amount of the total emissions that could be attributed to the steel producing plants in each county.

Data for the 41 steel-producing and the 114 non-steel counties of Ohio and Pennsylvania (the states with the largest number of steel-producing counties) were examined to discern possible patterns of relationships among variables for which values have been obtained on the data entry forms. The NEDS retrieval of emissions values from individual SIC 3312 point sources indicated that a number of counties with significant steel production operations were omitted from the earlier listing provided by EPA (for which the steel/non-steel county analysis was requested). The counties added for this two-state analysis were Belmont, Columbiana, Marion, Muskingum, and Tuscarawas in Ohio and Lancaster, Pennsylvania.

TABLE IV-33

RESPIRATORY DISEASE CATEGORIES FOR WHICH MEDICARE CLAIMANT RATES IN "STEEL-PRODUCING" COUNTIES ARE GREATER THAN RESPECTIVE STATE RATES

STATE	COUNTY/CITY	WHITE MALES CATEGORY						WHITE FEMALES CATEGORY					
		1	2	3	4	5	6*	1	2	3	4	5	6*
Alabama	Etowah					x							x
	Jefferson						x						
California	Alameda				x								
	Los Angeles											x	
	San Bernadino	x	x		x	x		x	x		x		
Colorado	Pueblo	x	x		x						x	x	
Connecticut	Fairfield	x	x				x	x	x	x			x
Delaware	New Castle												
Florida	Hillsborough	x	x	x	x	x	x	x	x	x	x		x
	Martin												
Georgia	Fulton												
Hawaii	Honolulu						x				x		x
Illinois	Cook												
	Madison	x	x	x	x	x	x	x	x	x	x	x	x
	Whiteside	x			x	x		x		x	x		
Indiana	Allen						x						
	Henry	x	x	x	x	x	x	x	x	x	x		
	Howard	x	x	x	x	x	x	x	x	x	x		
	Lake												
	Marion						x						
	Porter	x			x		x	x			x	x	
	Vigo	x	x		x	x	x	x		x		x	x
Kentucky	Boyd	x	x		x		x	x	x	x	x	x	x
	Campbell												x
	Daviess	x	x	x		x		x		x		x	
Louisiana	Tangipahos Parish	x	x	x	x	x	x	x	x	x	x	x	x
Maryland	Baltimore City+												
	Baltimore	x					x						x
Michigan	Macomb						x					x	x
	Wayne												
Minnesota	Ramsey	x										x	x
	St. Louis	x			x	x				x		x	x
Missouri	St. Louis City+												
	St. Louis												

RESPIRATORY DISEASE CATEGORIES FOR WHICH MEDICARE CLAIMANT RATES IN "STEEL-PRODUCING" COUNTIES ARE GREATER THAN RESPECTIVE STATE RATES

STATE	COUNTY/CITY	WHITE MALES CATEGORY						WHITE FEMALES CATEGORY					
		1	2	3	4	5	6*	1	2	3	4	5	6*
New Jersey	Middlesex	x	x	x	x	x	x				x	x	x
New York	Albany	x	x		x	x	x	x			x	x	
	Chautauqua	x	x	x	x	x		x	x	x	x	x	
	Erie	x			x	x				x		x	
North Carolina	Mecklenburg												
Ohio	Ashtabula		x		x				x				
	Belmont+												
	Butler	x	x	x		x		x		x	x		
	Columbiana+												
	Cuyahoga						x						
	Jefferson	x	x	x	x	x	x	x	x	x	x	x	x
	Lake	x	x		x		x			x			
	Lawrence	x	x	x	x	x	x	x	x	x	x	x	x
	Lorain	x	x			x				x		x	
	Lucas					x			x			x	x
	Mahoning				x						x	x	
	Marion+												
	Muskingum+												
	Richland		x	x		x		x	x	x		x	
	Scioto	x	x		x			x	x		x	x	x
	Stark	x			x	x	x	x	x		x	x	x
	Trumbull	x			x	x		x	x	x	x	x	x
	Tuscarawas+												
	Washington	x	x	x	x	x		x	x	x	x	x	x
Oklahoma	Tulsa						x						
Oregon	Multnomah				x	x							
	Yamhill		x				x				x		
Pennsylvania	Allegheny			x	x								x
	Beaver		x			x		x	x		x		x
	Berks						x						
	Bucks				x								x
	Butler	x	x	x	x	x		x		x		x	x
	Cambria	x	x	x	x		x	x	x	x	x		x
	Chester					x				x			
	Dauphin												
	Erie	x			x	x		x		x	x	x	
	Lancaster+												
	Lawrence	x	x		x	x		x	x	x	x	x	x
	Lebanon		x										
	Mercer	x	x	x	x	x		x	x	x	x		

TABLE IV-33 (continued)

RESPIRATORY DISEASE CATEGORIES FOR WHICH MEDICARE CLAIMANT RATES IN "STEEL-PRODUCING" COUNTIES ARE GREATER THAN RESPECTIVE STATE RATES

STATE	COUNTY/CITY	WHITE MALES CATEGORY*						WHITE FEMALES CATEGORY*					
		1	2	3	4	5	6	1	2	3	4	5	6
Pennsylvania (Continued)	Mifflin	x		x	x	x	x	x		x	x		
	Montgomery												
	Northampton				x								
	Northumberland	x	x			x	x		x				
	Philadelphia												x
	Venango	x	x	x	x	x		x	x	x	x	x	x
	Warren	x	x	x	x	x		x	x	x	x		
	Washington	x	x	x	x	x	x	x	x		x	x	x
	Westmoreland	x	x	x	x	x	x	x		x	x	x	
South Carolina	Darlington	x	x	x		x	x	x	x		x	x	x
	Georgetown	x	x	x	x		x			x			
	Richland											x	
Tennessee	Hamilton					x	x					x	x
	Knox						x						x
	Roane	x	x	x	x	x	x	x	x				x
Texas	Austin												x
	Chambers	x	x	x	x	x	x	x	x	x	x		x
	El Paso						x					x	x
	Gray	x	x	x	x	x	x	x	x	x	x	x	x
	Guadalupe					x							x
	Harris						x					x	x
	Morris	x	x	x	x	x	x	x	x	x		x	x
	Tarrant												
Utah	Utah	x	x	x	x	x		x	x	x	x	x	x
Virginia	Buchanan	x	x	x	x	x	x	x	x	x	x	x	x
Washington	King	x			x		x						x
West Virginia	Brooke		x										
	Cabell				x						x	x	
	Hancock	x	x		x		x	x	x		x	x	x
	Marion						x						
Wisconsin	Milwaukee												

* Category 1: All Respiratory Diseases (ICDA 460-519)
2: Acute Infection, Except Influenza (460-466)
3: Influenza (470-474)
4: Pneumonia (480-486)
5: Bronchitis/Emphysema/Asthma (490-493)
6: Other Respiratory Diseases (500-519)

+ Data not available at this time.

Cluster diagrams were prepared indicating observed relationships between (1) the Medicare hospitalization rate among white males for all respiratory diseases (ICDA 460-519) and monitored ambient TSP concentrations, (2) the Medicare hospitalization rate among white females for bronchitis/emphysema/asthma (ICDA 490-493) and TSP emissions, and (3) the Medicare hospitalization rate among white males for "other respiratory/bronchopulmonic disease" (ICDA 460-466, 500-519) and white-male age-adjusted mortality rate for the same disease classification.

The first two sets of cluster diagrams showed no significant associations for either the Ohio or Pennsylvania counties. The third set of cluster diagrams (Medicare hospitalization rates and age-adjusted mortality rates) for "other respiratory/bronchopulmonic disease" for white males showed the most pronounced correspondence. Correlation coefficients were computed separately for Ohio and Pennsylvania steel counties and were above 0.60 for Pennsylvania non-steel and Ohio steel counties. It was 0.21 for the Ohio non-steel counties.

d. Geographic Distribution of Morbidity--

The male and female morbidity rates (i.e., number of Medicare claimants hospitalized in 1971-1973) in each of six respiratory disease categories for the 155 Ohio and Pennsylvania counties were listed in rank order, and the respective ranks were entered onto the individual county data entry form. Ranks were also entered on "working maps" of Ohio and Pennsylvania showing the morbidity rates for ICDA categories 460-466 (acute infection, except influenza) and 490-493 (bronchitis/emphysema/asthma).

There was no consistent steel industry locational significance to the county morbidity rankings displayed on the maps. However, shortage of funds and time precluded meteorological modeling of any kind. These same constraints made it impossible to include other industries in the analysis.

In summary, no consistent relationship between steel production and morbidity rates in the same counties has been shown by this limited pilot investigation. The absence of such a relationship is believed to be extremely important to evaluations associating adverse human health effects, such as respiratory diseases, to airborne emissions from the steel industry alone or to a significant degree. Similarly, the absence of a strong steel industry-morbidity relationship demonstrates the need for refined industry-morbidity-mortality research to assist in guiding abatement and control research priorities on the one hand, and establishing regulation and environmental protection priorities on the other.

8. Computer-Prepared Maps Comparing Pollution/Morbidity Factors

To demonstrate the relevance to health effects analyses of geographical relationships derivable from the Medicare morbidity tapes, several different procedures were explored. These procedures feature combining statistics with maps of the United States that show data and geography to the county level of detail. Some maps were prepared entirely by computer, utilizing techniques, programs, and equipment readily available to Health Effects Research Laboratory. Others were prepared by "hand," such as the preceding Ohio and Pennsylvania maps which identify respiratory disease rates of steel counties by encircling the data.

For the computer-prepared maps, EPA Research Triangle Park equipment was used exclusively. The maps can display data from Medicare morbidity files and many other sources, relatively inexpensively, in multi-color format, and without significant delays or additional specialized programming expenditures.

Display of statistical data in map format is helpful analytically in revealing efficiently the presence, and absence, of relationships that are difficult to detect from tabulations alone or without detailed familiarity with geopolitical spatial relationships. For health effects analyses, it is important to note that computer techniques enable designation of one or two variables, or many variables. The variables are not limited to the Medicare morbidity tapes, but can be selected from EPA and many other data files.

Two maps were prepared, utilizing EPA facilities, to assist in the analysis of possible relationships between respiratory diseases (as quantified by the Medicare morbidity tapes) and steel-producing facilities. Respiratory diseases were selected as indicative of adverse health effects caused by industrial airborne emissions of concern to EPA.

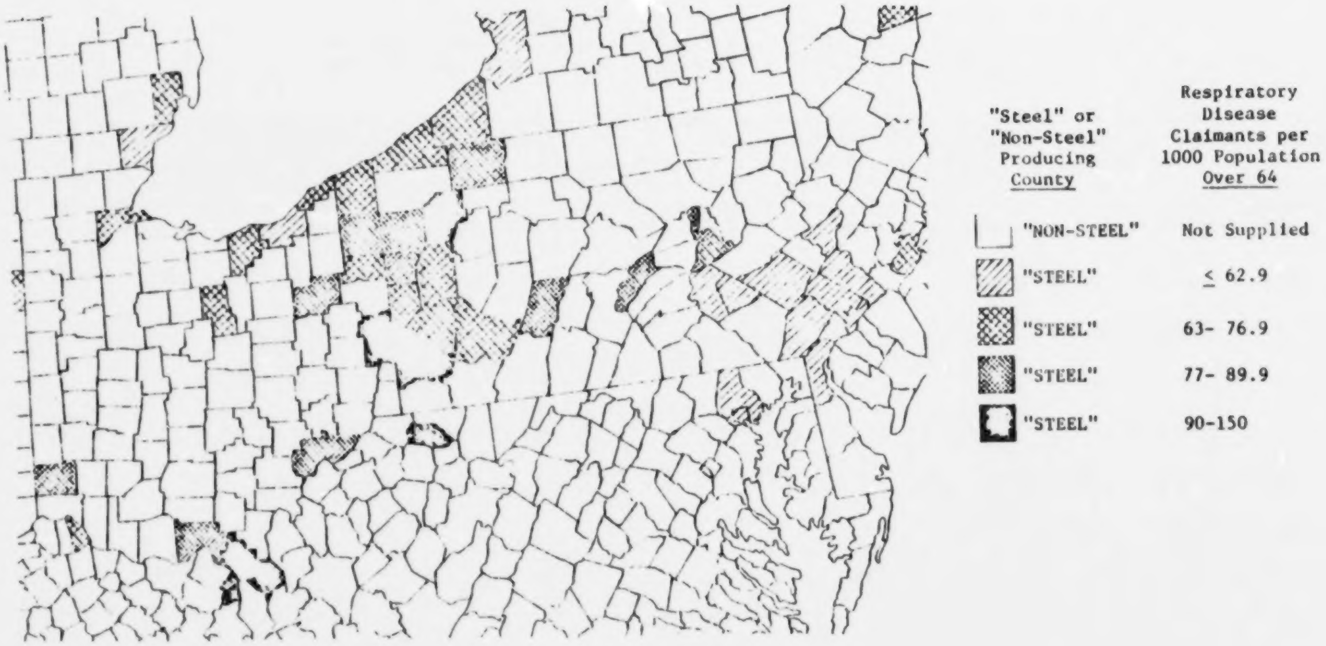
The first map, Fig. IV-8, is a cut-out of a computer-prepared U.S. map identifying counties in Ohio and Pennsylvania and adjacent areas listed in the NEDS data base as containing one or more SIC 3312 steel-producing establishments. These "steel" counties are then sub-classified graphically, depicting one of four categories of respiratory disease hospital claimants rates per 1000 white males over 64 years of age. The respiratory disease data are taken from the Medicare morbidity tapes; the population data from POPATRISK. It will be recalled that county-by-county comparisons with state average rates for six different respiratory disease categories are shown in Table IV-33.

The second computer-prepared map, Fig. IV-9, illustrates the use of this technique in analyzing possible "steel production" and respiratory disease relationships in greater detail. In this more complex map, the hospital respiratory disease rates for the counties that do not produce steel are shown in blue.

This more complex map also differentiates among the steel-producing counties in terms of the estimated TSP emissions from the steel industry as a percentage of all point sources in the county. In Ohio and Pennsylvania there are a total of 21 counties with steel plant estimated TSP emissions aggregating up to 12.7 percent of the county's point source TSP emission. These 21 counties are shown in red.

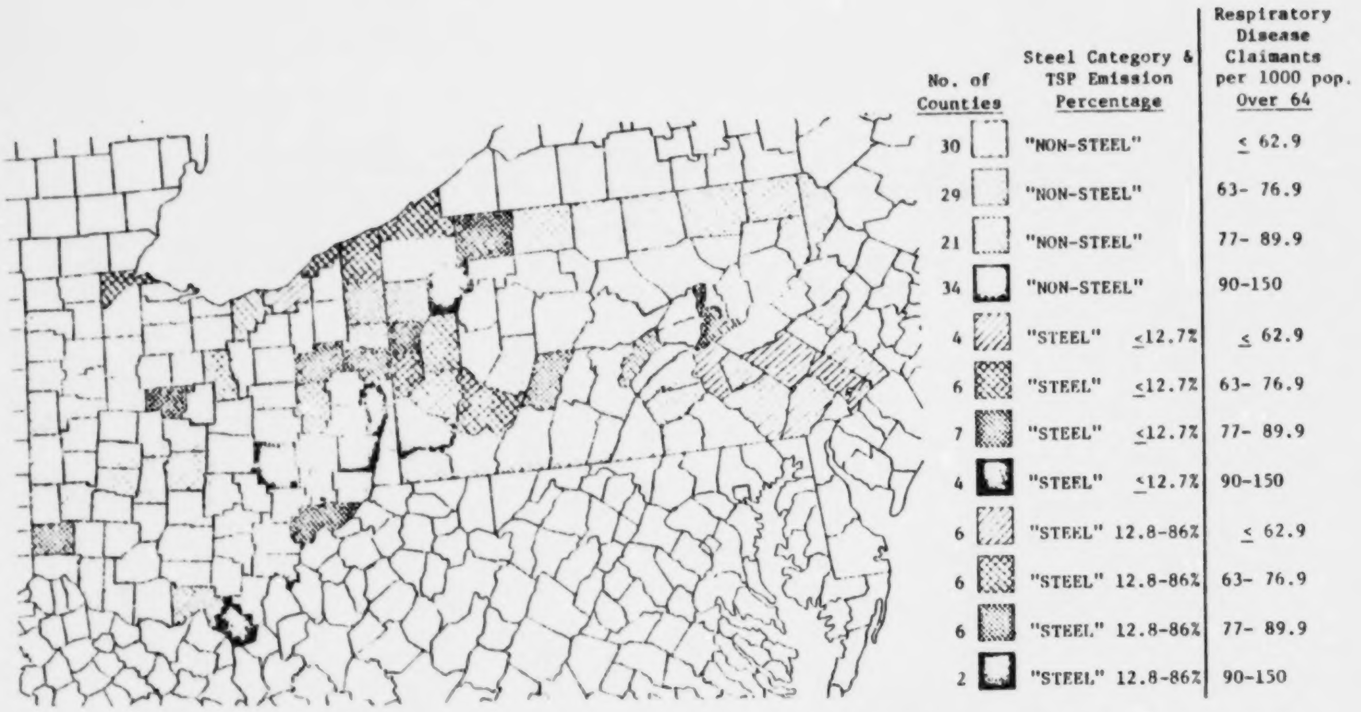
The 20 remaining steel-producing counties have steel-producing sources that emit from 12.8 percent to 86 percent of the TSP point source emissions in each county. These 20 counties are shown in black.

For each of these three categories of counties (no steel production, counties with steel industry sources aggregating 12.7 percent or less of all TSP emissions, and counties with steel-producing sources emitting more than 12.7 percent of all TSP emissions), the Medicare respiratory disease hospitalization patient rates are shown by one of four categories of rates per 1,000 white males over 64 years of age.


"Steel" or "Non-Steel" Producing County
Respiratory Disease Claimants per 1000 Population Over 64
"NON-STEEL" Not Supplied
"STEEL" ≤ 62.9
"STEEL" 63- 76.9
"STEEL" 77- 89.9
"STEEL" 90-150

[Illustration of computerized multi-color mapping capabilities (up to 29 color/line categories, variable size-scaling) of EPA's County Shading Program, using UNIVAC 1110 computer system and CalComp pen plotter]

Fig. IV-8. Steel-producing counties categorized by Medicare respiratory disease hospital patients per 1000 population over 64 years of age, 1971-1973. Extract of U.S. map. "Non-steel" counties are blank.


No. of Counties
Steel Category & TSP Emission Percentage
Respiratory Disease Claimants per 1000 pop. Over 64
30 "NON-STEEL" ≤ 62.9
29 "NON-STEEL" 63- 76.9
21 "NON-STEEL" 77- 89.9
34 "NON-STEEL" 90-150
4 "STEEL" ≤12.7% ≤ 62.9
6 "STEEL" ≤12.7% 63- 76.9
7 "STEEL" ≤12.7% 77- 89.9
4 "STEEL" ≤12.7% 90-150
6 "STEEL" 12.8-86% ≤ 62.9
6 "STEEL" 12.8-86% 63- 76.9
6 "STEEL" 12.8-86% 77- 89.9
2 "STEEL" 12.8-86% 90-150

[Illustration of computerized multi-color mapping capabilities (up to 29 color/line categories, variable size-scaling) on Government-furnished equipment used by SSI. This computer map has been assembled with data from National Emissions Data System (EPA), Census, and Health Care Finance Administration to illustrate flexibility of this graphic system in an HERL context. Air quality, by pollutant, fuels consumption, and other parameters may also be displayed. The data are presented only as illustrative of a technique, and not as a finished analysis.] NOTE: Original of this figure was produced in three colors by the County Shading Program implemented on EPA's UNIVAC 1110 computer system using a CalComp pen plotter.

Fig. IV-9. Ohio and Pennsylvania counties classified by percentage of steel industry suspended particulate emissions, and Medicare claimant respiratory disease hospital rates per 1000 population over 64 years of age.

As a partial analysis of the co-locational significance of steel-producing facilities to a county's Medicare respiratory disease rates and experience, it is interesting to compare the counties that have the highest respiratory disease rates.

Steel-producing counties--Only two of the 20 "black" counties (the higher percentage groups of steel industry emissions) are also in the highest category of respiratory disease rates. The other 18 counties are equally distributed among the three remaining classes of respiratory disease rates. These data do not show a direct association between steel industry TSP emissions and higher respiratory disease rates for the co-located county population.

Among the 21 "red" counties (the lower percentage group of steel industry emissions), only four counties are in the highest category of respiratory disease rates. An equal number of counties is in the lowest category of respiratory disease rates. Of the remaining 13 counties, seven are in the next highest category of respiratory disease rates. No clear conclusion can be drawn from this group of steel producing counties and their respiratory disease experience with the population of Medicare age.

Non-steel-producing counties--The health status of the population in non-steel-producing counties, if measured in Medicare respiratory disease rates, is lower than that of the steel-producing counties. This generalization applies both to the "steel" counties with a low percentage of emissions from steel producers, and to the counties in the higher category of steel production emission sources.

Of the "non-steel" counties, 34 had the highest of the four respiratory disease rate classes; 21 fell in the next highest category; 29 in the next lower category, and the remaining 30 counties had the lowest category of Medicare respiratory disease cases.

Incomplete analysis of the non-steel-producing counties, in terms of Medicare respiratory disease rates, indicates there are many sources of emissions, types of emissions and other factors, that must be taken into consideration in comparing steel and non-steel counties. Many industries and natural events contribute to air pollution in specific localities. It is quite possible that a detailed analysis would disclose hazardous sources within certain non-steel counties that would explain their higher respiratory disease rates. Time and budget did not permit an analysis of possible pollutant sources in the non-steel-producing counties, or of other factors to be considered in a thorough analysis. Besides meteorology, to account for emissions from upwind sources, other pertinent factors are population relocations of the adversely affected, and death rates by age for both steel and non-steel counties.

The respiratory disease morbidity rates of the steel- and non-steel-producing counties raise strategic, tactical, and technical questions for achieving health-improvement or health-protection goals by reducing emissions on a selected industry basis. Many of these questions may be answered by analyses of trends in respiratory diseases compared with estimated emissions and technically reliable monitoring of the ambient air. On the other hand, if reduced emissions are not being reflected in reduced respiratory disease, then an entirely new set of questions arises as to the validity of emissions-reducing programs and associated billions of dollars of expenditures that have been justified on health protection and improvement grounds.

SECTION V

REVIEW AND DISCUSSION

This project featured "non-confidential" portions of the computerized Medicare hospitalization 20 percent sample file containing neoplasm, respiratory, and digestive disease diagnoses for the years 1971, 1972, and 1973. Extracts in tape form were received from Social Security Administration, the records were sorted and otherwise reformatted for convenient use on EPA's Research Triangle Park computer, retrievals of pertinent data were programmed, and illustrative presentation of the results made. Other data resources supplied information on county industrial development, demography, air quality, emissions, and indicators of health status.

The geographic focus of the study varied according to task, with most attention being given to three areas in Florida and Pennsylvania. For some aspects of the work, statistics were compiled nationwide. A special analysis compared counties with and without steel-production facilities in terms of respiratory diseases, and with county-by-county details in Ohio and Pennsylvania. At several junctures, data from counties with air pollution episode or other relevant aspects were given consideration and appropriate data incorporated.

Since the objective of this work was to demonstrate the potential utility of the Medicare hospitalization file and not to conduct a definitive analysis of disease patterns, little time was available with which to compile comprehensive environmental statistics for all the counties and to prepare analyses in depth. Rather, data sets were chosen to feature and demonstrate methodologies for applying the Medicare morbidity data in the context of interests of primary concern to Health Effects Research Laboratory.

The credibility/reliability of the Medicare files was evaluated by comparisons with National Cancer Institute "survival times" for patients 65 years of age and over. At the national level of aggregation, and with qualifications described in the text, the Medicare record contents are very closely comparable with NCI record contents. The Medicare data may in fact be more representative of the national population over age 64 than NCI data, which are taken from two statewide tumor registries and about a dozen metropolitan areas aggregating about 10 percent of the total United States population. The Medicare records are from all U.S. counties; the sample includes five million of the 25 million covered under the Medicare and Railroad Retirement Act insurance procedures.

From a health effects perspective, it is believed necessary that national control and abatement programs of EPA justified to protect human health be quantifiable systematically in human health terms. No alternative source is available and superior to that of the Medicare data files and their potential.

Major project activities were discussed in Section IV. Results of various computer-assisted analyses are presented there, and summarized below.

A. BERKS COUNTY--MORBIDITY/MORTALITY COMPARISON BY YEAR

The retrievals by age, sex, race and time of hospitalization such as demonstrated for residents of Berks County in Section IV, give a clear indication of the power of the Medicare data to measure accumulating environmentally related diseases with a comparatively long latency and time-to-death period (cancer) in a county's resident population. A procedure for estimating the net change in numbers of cases with selected cancer diagnoses was demonstrated for each of the years 1971, 1972, 1973. The kind of "running inventory" of individuals with diseases of special environmental interest demonstrated in Section IV is superior in certain respects to mortality statistics and in all cases is a unique, valuable indicator of environmental health in local areas.

Mortality descriptors alone have many deficiencies as indicators of the human health status with respect to the environment. The most prominent of these statistical deficiencies are the long time lag between exposure and death, and the high probability that death will occur from causes unrelated to unhealthy environmental conditions. Morbidity descriptors, on the other hand, are more responsive as human health indicators to changes in the environment--whether good or ill. The mortality and morbidity files should be used in mutual support. Air quality improvements can be expected to improve health while degradation of the air can be expected to have the opposite effect. By comparing ambient measures, emissions, and morbidity/mortality experiences of counties, metropolitan areas, and other regions, EPA control and abatement priorities based on human health considerations may also be quantified systematically from both hospitalization and "running inventory" of survivors of selected environmentally related diseases.

Distribution of cases by age at the time of disease identification, or at any other specified time as is feasible from Medicare records, offers an outstanding opportunity to note significant increases in disease occurrence levels which may be caused by the introduction of a hazardous condition in a particular locality. The Medicare files facilitate detailed comparisons among localities with respect to time of disease diagnosis, and patient age/sex/race. If there has been a definitive and time-specific change in county environmental quality and a disease onset is expected from prolonged exposure to the conditions introduced, the procedures demonstrated with the Berks County Medicare data may be useful in determining the latency period for the disease and confirming/denying the association with the condition being evaluated. No comparable timely opportunity is offered by mortality data.

Longitudinally assembled admissions/diagnoses data, in conjunction with the survival-time statistics illustrated in Section IV can be of great value in monitoring environmental health in terms of cases being added to or subtracted from the county-resident environmentally-related disease pool. Considerable potential is seen in the richness of morbidity-related information available in the Medicare files which is not available through mortality data based only on underlying cause. This additional detail would be of use in studying health/environment histories of both deceased and surviving individuals.

Tabulations of new cases conceptually assembled routinely from Medicare records, when combined with knowledge of the time-spans and development sequences between first manifestations of diseases and later death, can serve to provide EPA with early indicators of (a) factors in the county or region which may pose increased health risks and (b) the effectiveness of control measures in removing health-threatening factors.

In considering potential uses of the Medicare files for the types of analysis demonstrated in Berks County, the limitations of the data must be kept in mind. Being a sample, the Medicare claimant records may not be representative of the true prevalence of the diseases reported, particularly in areas with small numbers of residents of over relatively few years. The medical histories are incomplete; most claimants become eligible for Medicare at age 65, only in-patient hospital treatments are reported, and deaths which do not occur during a hospital stay are not indicated. The Continuous Work History Sample (CWHS) of the Social Security Administration offers some remedy for these shortcomings, as the CWHS reports all deaths in its sample. Time, work scope and budget did not allow this CWHS computerized file to be utilized.

B. DAILY MORBIDITY/MORTALITY DURING PITTSBURGH EPISODE OF 1973

The daily morbidity statistics compiled for Allegheny County and the other counties of the Pittsburgh Standard Metropolitan Statistical Area are of special interest to the EPA's Health Effects Research Laboratory as supplements to mortality data, and are particularly relevant to EPA's continuing investigation of "excess deaths" during and immediately following a pollution episode such as that of 1973. Because hospital admissions/discharges for a disease do not have the unique conclusiveness which characterize the event of death (and attributed underlying cause), three different classifications of morbidity were used in the 1971, 1972, and 1973 retrievals of Medicare hospitalization data. These enumerated (1) all admissions during the episode's 13 days and the 13 days before and after the episode, (2) only the individual's first admission for a disease classification during the time period, and (3) the number of individuals admitted during each 13-day sub-interval.

Analysis of this Medicare data discloses that the respiratory disease hospitalization counts were not above the three-year average each day of the 1973 episode, as was true for the "all-causes" mortality data. From the perspective of environmental linkages exclusively, there were significantly more respiratory disease cases occurring during the 13-day episode period than in the other eight 13-day 1971-1973 reference intervals. The occurrence of four deaths from respiratory disease following admission during the episode was in striking contrast to the one or two hospital deaths recorded for each of the previous two years but the small numbers do not justify conclusions. For the Pittsburgh episode, the additional respiratory disease hospitalizations, immediately prior to the episode and during the episode, appear to provide a clear environmental health justification for the episode declaration and accompanying precautionary measures. It would be interesting to extend the Pittsburgh analysis to other areas experiencing episode or near-episode conditions to determine if they also experienced abnormally high respiratory disease hospitalizations.

The morbidity data offer two major advantages over the mortality data for studies of changes in daily rates. First, the morbidity data include non-fatal disease conditions, and, therefore, are not constrained to the stringent requirement of death. This means higher daily counts of cases, a greater variety of cases, a more complete understanding of environmental health effects, and the practical feasibility of obtaining more definitive information from the patient. Second, the date of hospitalization on the Medicare record is generally the same as, or shortly after, the date of environmental interest, while the date of death for mortality cases has a more variable and uncertain relationship to the date of fatal illness manifestation and to environmental dates of interest.

The daily morbidity data provided by the Medicare files may be of particularly great value in the process of pollutant standard-setting and emergency abatement procedures. The data may be used (as in Pittsburgh) to demonstrate relationships between air quality and human hospitalizations for certain diseases. It may also be possible to quantify emissions/ambient/disease correspondence, [illegible] to specify pollutant level(s) which precede(s) a significant increase in [illegible]tal admissions.

No other source of morbidity data is known which can so readily provide the detail available through the Medicare files for any county or aggregates of counties, including the daily, monthly or other time period of hospitalization. If more geographic or disease detail is needed, it is believed quite feasible to expand the present 20 percent sample to 40 percent, or 80 percent or higher sample. The additional costs are believed relatively minor, as data on all Medicare hospitalizations are routinely required for reimbursements, no new steps are involved, and expanding the sample only involves additional ICD encoding and related computerization steps.

C. MORBIDITY IN STEEL-PRODUCING AREAS

The comparison of morbidity rates in counties with steel-production establishments versus morbidity rates in non-steel counties across the United States was designed to show how the various component data resources can be brought together for analytical purposes. The steel industry was selected because of EPA's special interest. Other industries can also be selected, as well as groups of industries. Similarly, counties and their associated morbidity/mortality may be classified and analyzed as rural-urban, including estimated emission categories, ambient categories, summer-winter temperature, fuels consumption, and so forth.

This steel-producing and non-steel-producing counties comparison, described in Section IV, has demonstrated a number of informative ways by which the Medicare computerized files help determine whether the operations of a specific industry pose health risks to the surrounding resident population. Methodological refinements can help the Medicare data to be used even more effectively for this kind of epidemiological research function.

The fact that the hospitalization rates for respiratory disease were generally lower in the steel-producing counties than elsewhere, and showed no clear relationship between the relative importance of steel-producing sources of TSP and respiratory diseases runs counter to expectation. Common sense suggests that steel industry emissions are not healthful. Further, adverse health effects from high TSP levels and from exposure to steel production processes have been documented. For example, during the Allegheny County episode discussed above, respiratory disease hospitalizations increased, in conformance with expectations.

A number of factors affect the validity of findings based on Medicare hospital admissions in counties with a particular industry. First, the Medicare sample, predominantly individuals over 65, has not been demonstrated to represent the health effects of younger age groups to a polluted environment. Some diseases claim their victims at an earlier age, and higher pollution levels may have exacted conclusive effects before the victims reached the Medicare age. Hospital admissions for Medicare residents of industrialized and more polluted areas may also be lower than expected because the more susceptible individuals may have relocated to more healthful environments. With increasing age, many other disease/accidents present important competing risks, thus masking the full health impact of environmental pollution in heavily industrialized areas.

The use of county level of detail for the study of morbidity and mortality data poses a problem in examining the contributions of point-specific sources such as steel mills. The industrial facility should, at the very least, be determined to be a major emitter of a pollutant within the county. In the present project, all steel-producing counties were considered and although the relative importance of steel industry sources was determined, the significance of the variations was not analyzed. Time and budget constraints also limited the consideration given to the presence of other types of industrial establishments. With the EPA/NEDS data base and the Census/SSA files of four-digit SIC employment, these latter difficulties can be overcome.

Finally, the scope of this project did not permit examination of some important variables; more thorough investigation of all possible correspondences would be required in a definitive study. Inclusion of additional variables (ambient measures, estimated emissions, and trends for steel and non-steel producing counties) should increase the health-hazard-detection sensitivity of the methods already described and suggest new pollution abatement strategies based on human health protection. Reduction of emissions across-the-board for a particular industry, such as steel, represents one of many alternative approaches with advantages and disadvantages to be considered carefully before implementation.

In summary, the comparatively lower respiratory disease experience of the steel producing counties suggests the necessity for further study, for if confirming/verifying results are obtained, serious questions are raised concerning acceptable background levels, abatement and control of the particles (size, composition) that are harmful and abatement/control policy trade-offs between health purposes and aesthetic purposes.

D. MIGRATION OF MEDICARE CLAIMANTS

The examination of Medicare claimant migration (Section IV) did not indicate as widespread a relocation pattern as expected from the retirement age population, especially in the high in-migration counties of Florida. Extremely low rates of county/state relocations were observed in the Medicare sample study areas of Florida and Pennsylvania; once hospitalized, changes in residency are few. It may be that most individuals who change residences do so shortly after the retirement date, or while still in relatively good health. The Medicare records alone are, therefore, not a good source of data either on migration patterns or on the contribution of migration to health effects of various geographic regions, but must be used in conjunction with other sources.

A more complete representation of county/state environmental exposures prior to age 65 can be obtained from longitudinal employment and location history computerized files now being obtained from Social Security Administration. These files may be linked with the Medicare morbidity files to show both industry and location during pre-retirement years and subsequent relocations. With this linkage, long latency period diseases may be identified and analyzed in conjunction with environmental exposure areas.

E. FINDINGS AND CONCLUSIONS

The Medicare hospital discharge survey sample is a source of morbidity data with unique value for disclosing and monitoring relationships between the environment and the health of an important segment of the population. It is a growing data base of medical records, is routinely collected, covers all parts of the country, is maintained in "computerized" format, and the files may be assembled longitudinally to disclose the complete Medicare hospitalization history of all sample selectees, regardless of disease or place of hospitalization. These files provide daily information on primary causes of each hospital admission geographically identified by the claimant's county of residence. The sex, race and age of the claimant at the time of admission are also shown.

The data from the Medicare file can be utilized to estimate national, regional, state, or county totals of hospital admissions by discrete types of diseases. In assessing the credibility of the Medicare data, a close correspondence was shown to exist with other morbidity/mortality surveys and files.

Procedures, with tables, were demonstrated which measure for each year the addition to and loss from the pool of residents having specified diseases. The capability for estimating the population that has experienced one or more hospitalizations, by ICD code, for each of the 3,050 counties, has many environmental health applications. It provides an additional perspective to raw admissions counts, or counts of mortalities by cause. The scope of these environmental health applications could not be addressed in this pilot project, but the procedure demonstrated provides an additional tool for comparing counties and metropolitan areas by population burden of environmentally-related diseases. Comparisons can be sharpened by utilizing ambient and emissions descriptors for the areas being analyzed to differentiate, if possible, between (a) "normal" or "background" case-load levels of the diseases of particular interest; (b) additional case-loads that might be attributed to chronic low-level exposures to pollutants; and (c) higher case-loads caused by acute or emergency episodic exposures to pollutants. Daily counts were taken which noted the peaking effect of hospital admissions during the pollution episode in Allegheny County, Pennsylvania, 1973.

A preliminary study of a single industry (steel) was performed, demonstrating how the files may be used in conjunction with other EPA computerized files to assess the relationship between selected diseases and the presence or absence of specific industrial sources of emissions.

The extent of migration exhibited in Medicare claimant hospitalization histories was investigated with the finding of relatively few changes in residence after a hospitalization occurs.

The pilot study demonstrations indicate the flexibility of the Medicare data for a variety of purposes and suggest that, with appropriate supplementary refinements, this resource may be developed to have important uses in the Health Effects Research Laboratory and elsewhere in EPA. As examples, the Medicare files offer assistance in emergency episode determination and conclusion decisions, differentiating between industry and non-industry sources of environmental diseases, in focusing regulation and enforcement resources on factors most injurious to human health, and other evaluations in EPA that directly or indirectly are concerned with human health and the environment.